ORIGINAL RESEARCH

Ecology and Evolution

WILEY

Adaptive landscape genetics and malaria across divergent island bird populations

Claire Armstrong[1] (iD) | **Richard G. Davies[1]** | **Catalina González-Quevedo[1,2]** (iD) | **Molly Dunne[1]** (iD) | **Lewis G. Spurgin[1]** (iD) | **David S. Richardson[1]** (iD)

[1]School of Biological Sciences, University of East Anglia, Norwich, UK

[2]Grupo Ecología y Evolución de Vertebrados, Instituto de Biología, Facultad de Ciencias Exactas y Naturales, Universidad de Antioquia, Medellín, Colombia

Correspondence
David S. Richardson, School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK.
Email: David.Richardson@uea.ac.uk

Abstract

Environmental conditions play a major role in shaping the spatial distributions of pathogens, which in turn can drive local adaptation and divergence in host genetic diversity. Haemosporidians, such as *Plasmodium* (malaria), are a strong selective force, impacting survival and fitness of hosts, with geographic distributions largely determined by habitat suitability for their insect vectors. Here, we have tested whether patterns of fine-scale local adaptation to malaria are replicated across discrete, ecologically differing island populations of Berthelot's pipits *Anthus berthelotii*. We sequenced TLR4, an innate immunity gene that is potentially under positive selection in Berthelot's pipits, and two SNPs previously identified as being associated with malaria infection in a genome-wide association study (GWAS) in Berthelot's pipits in the Canary Islands. We determined the environmental predictors of malaria infection, using these to estimate variation in malaria risk on Porto Santo, and found some congruence with previously identified environmental risk factors on Tenerife. We also found a negative association between malaria infection and a TLR4 variant in Tenerife. In contrast, one of the GWAS SNPs showed an association with malaria risk in Porto Santo, but in the opposite direction to that found in the Canary Islands GWAS. Together, these findings suggest that disease-driven local adaptation may be an important factor in shaping variation among island populations.

KEYWORDS

birds, genome-wide association studies, Haemosporidians, landscape genetics, malaria, pathogen-mediated selection, Toll-like receptor

1 | INTRODUCTION

Spatial variation in the environment is a key force driving local adaptation and population divergence (Hereford, 2009). There is increasing evidence that fine-scale changes in environmental conditions can result in variation in selection pressures over small geographic distances (Garroway et al., 2013; Langin et al., 2015; Richardson, Urban, Bolnick, & Skelly, 2014) and that fine-scale adaptation can persist despite the homogenizing effects of gene flow (e.g., Lenormand, 2002). Adaptation to spatially heterogeneous environmental conditions can also facilitate balancing selection, where genetic variation is maintained within and between populations due to differential selection on genetic variants (Bockelmann, Reusch, Bijlsma, & Bakker, 2003; Levene, 1953; Schmidt, Bertness, & Rand, 2000).

Pathogens are major drivers of evolution, exerting strong selective pressures on their hosts (Fumagalli et al., 2011). The high genetic variation found at many genes involved in immune processes is thought to be maintained by pathogen-mediated balancing selection (Bernatchez & Landry, 2003; Ferrer-Admetlla et al., 2008; Spurgin & Richardson, 2010). Spatial variation in pathogen-mediated selection has the potential to drive fine-scale heterogeneity in immunogenetic diversity (Larson, Lisi, Seeb, Seeb, & Schindler, 2016; Tschirren, Andersson, Scherman, Westerdahl, & Råberg, 2011), highlighting the importance of spatial scale in understanding pathogen-mediated selection.

Haemosporidians in the genera *Plasmodium*, *Haemoproteus*, and *Leucocytozoon* (hereafter termed malaria for simplicity) are protozoan parasites that infect the red blood cells of mammals, reptiles, and birds (Martinsen, Perkins, & Schall, 2008). Infection by malaria has been associated with increased mortality, decreased body condition, and reductions in fitness (Guggisberg, Sayler, Wisely, & Odom John, 2018; Knowles, Palinauskas, & Sheldon, 2010; Marzal, Bensch, Reviriego, Balbontin, & De Lope, 2008). The selective pressure exerted by malaria has driven the evolution of increased host resistance and tolerance (Atkinson, Saili, Utzurrum, & Jarvi, 2013; Hill et al., 1991), with evidence of local adaptation to spatially heterogeneous selection pressures (Loiseau et al., 2011; Piel et al., 2010). Malaria parasites are dependent on vector transmission to complete their lifecycle, and their spatial distributions are therefore constrained by the environmental niches of their insect vectors. Temperature, rainfall, and altitude play especially important roles in determining malaria prevalence (Illera, López, García-Padilla, & Moreno, 2017; Jones, Cheviron, & Carling, 2013; Padilla, Illera, González-Quevedo, Villalba, & Richardson, 2017). Water is essential for aquatic larval development of vectors, and topographic features that increase surface water persistence promote increased abundance (Ferraguti et al., 2016; Ganser et al., 2016; González-Quevedo, Davies, & Richardson, 2014). In addition, anthropogenic factors such as habitat degradation, agriculture, and urbanization can all influence malaria dynamics (González-Quevedo et al., 2014; Turcotte, Bélisle, Pelletier, & Garant, 2018; Yanoviak, Paredes, Lounibos, & Weaver, 2006). Together, these factors have the potential to shape fine-scale spatial structuring in pathogen selection pressures and host immunogenetic variation.

Studying pathogen-mediated selection has largely involved a candidate gene approach, where variation at genes with known, or predicted, host immunity function is investigated in relation to infection (Bernatchez & Landry, 2003; Netea, Wijmenga, & O'Neill, 2012). Many studies have focused on the major histocompatibility complex (MHC), a gene family that plays a key role in pathogen recognition in the adaptive immune system. However, a greater proportion of phenotypic variance in malaria response has been attributed to non-MHC genes (Jepson et al., 1997). Within the innate immune system (a first line of defense against infection), Toll-like receptors (TLRs) are a family of pattern-recognition receptors which have been linked to malaria resistance (Ferwerda et al., 2007; Mockenhaupt et al., 2006), and show evidence of pathogen-mediated balancing

selection (Ferrer-Admetlla et al., 2008; Fisher et al., 2011; Gavan, Oliver, Douglas, & Piertney, 2015). TLRs therefore represent important candidates for investigating the role of pathogens in maintaining host genetic variation.

An alternative to the candidate gene approach is the use of genome-wide association studies (GWAS). These enable detection of single nucleotide polymorphisms (SNPs) throughout the genome that show statistical associations with pathogen infection. In addition to identifying associations at known immune loci (Fellay et al., 2007; He et al., 2015; Wong et al., 2010), GWAS approaches may reveal novel candidate genes (Fu et al., 2012; Ravenhall et al., 2018; Thye et al., 2010) for further study of the evolutionary dynamics between host and pathogen.

Islands are excellent environments for investigating pathogen-mediated selection. In line with island biogeography theory (MacArthur & Wilson, 1967), pathogen diversity and abundance are often lower on islands compared with the mainland, simplifying the study of host–pathogen interactions (Pérez-Rodríguez, Ramírez, Richardson, & Pérez-Tris, 2013; Clark, Clegg, & Lima, 2014; but see Illera, Fernández-Álvarez, Hernández-Flores, & Foronda, 2015). Pathogen communities on each island are shaped by chance colonization and extinction events, which can result in distinct pathogen assemblages and selection pressures between islands (Fallon, Bermingham, & Ricklefs, 2005; Olsson-Pons, Clark, Ishtiaq, & Clegg, 2015; Wang et al., 2017). Limited gene flow in and out of islands also allows for stable communities of hosts and pathogens (Spurgin, Illera, Padilla, & Richardson, 2012), which may facilitate strong coevolutionary relationships.

Berthelot's pipit *Anthus berthelotii* is a small sedentary passerine endemic to three Macaronesian archipelagos (Figure 1). Following the colonization of the Madeiran archipelago from the Canary Islands ca. 8,500 years ago (Spurgin, Illera, Jorgensen, Dawson, & Richardson, 2014; but see Valente et al., 2017), there has been a lack of gene flow between the archipelagos (Illera, Emerson, & Richardson, 2007; Spurgin et al., 2014), potentially facilitating local adaptation and divergent selection (Armstrong et al., 2018). Malaria infection shows high spatial variability in this species, both
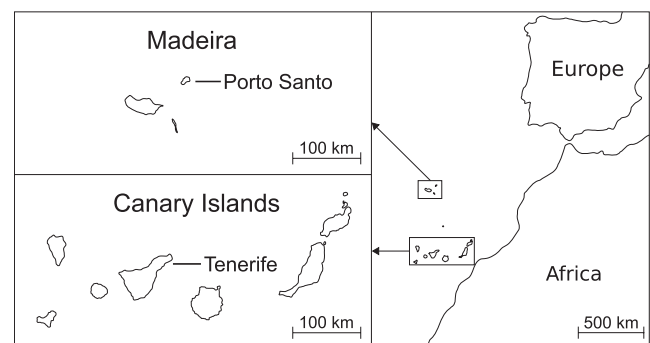
FIGURE 1 Map of Berthelot's pipit populations. Berthelot's pipits are found across all islands within the Madeiran archipelago (top left panel), the Canary Islands (bottom left panel), and the Selvagens archipelago, situated between the Canary Islands and Madeira

between and within islands, making it a highly suitable model for investigating the role of spatial scale in pathogen-mediated selection. Characterization of malaria throughout Berthelot's pipit populations (Illera, Emerson, & Richardson, 2008; Spurgin et al., 2012) found the highest prevalence of *Plasmodium* and *Leucocytozoon* infection on Porto Santo, whereas no infection was detected elsewhere in the Madeiran archipelago. Prevalence of malaria on Tenerife is influenced by a combination of climatic and anthropogenic effects (González-Quevedo et al., 2014; Padilla et al., 2017), with malaria undetected at high altitude. Associations between the distribution of MHC variants and environmental predictors of malaria infection have been detected (González-Quevedo, Davies, Phillips, Spurgin, & Richardson, 2016). No evidence of *Haemoproteus* infection has been found in this species.

Here, we test for associations between fine-scale patterns of genetic variation and malaria in Berthelot's pipits across two divergent populations (Tenerife and Porto Santo) to investigate the spatial scale of local adaptation in the presence of gene flow within a population. This study also allows us to test the repeatability of patterns of association across populations. These two islands, situated on different archipelagos, show high genetic divergence at neutral loci, with limited to no gene flow between them (Armstrong et al., 2018; Spurgin et al., 2014). Despite a sharp decline in overall genetic diversity associated with the initial colonization of Madeira (Armstrong et al., 2018), higher levels of TLR4 allelic and amino acid richness exist in Madeira compared to the Canary Islands (González-Quevedo, Spurgin, Illera, & Richardson, 2015). Furthermore, evidence of positive selection at TLR4 in Berthelot's pipits suggests it may be an evolutionarily important locus (González-Quevedo et al., 2015). In this study, we (a) test for associations between *Plasmodium* infection status (the only Haemosporidian genus commonly detected; >1% prevalence) in this species (Spurgin et al., 2012), and variation at TLR4 and two SNPs previously identified in a GWAS of malaria infection across Berthelot's pipit populations (Armstrong et al., 2018); (b) determine the environmental predictors of malaria risk on Porto Santo; (c) compare genetic associations with malaria risk in Porto Santo and Tenerife, utilizing the above measures of malaria risk for Porto Santo, and those previously calculated for Tenerife (González-Quevedo et al., 2016, 2014).

2 | MATERIALS AND METHODS

2.1 | Sample collection

Berthelot's pipits were sampled on Tenerife (Feb–April 2006, Jan–Aug 2009, April–May 2010, Jan–May 2011) and Porto Santo (Sept 2006, March 2009, April–June 2016). For samples collected prior to 2011, 30–96 birds were sampled widely across each island per season (Illera et al., 2007; Spurgin et al., 2012). For samples collected in Tenerife in 2011, attempts were made to catch one pipit in each 1 km^2 of suitable habitat, with a total of 388 birds sampled (González-Quevedo et al., 2014). On Porto Santo in 2016, all areas

of suitable and accessible pipit habitat throughout the island were surveyed. We attempted to catch every pipit encountered, resulting in a sample of 129 birds. In total, 780 birds were sampled across all years and islands.

Birds were caught in spring traps baited with *Tenebrio molitor* larvae. Each bird was fitted with a colored plastic ring or a numbered aluminum ring issued by the Spanish or Portuguese authorities as appropriate, to avoid resampling individuals. Blood samples (ca. 50 µl) were collected by brachial venipuncture and stored in 100% ethanol in a screw-top Eppendorf tube at room temperature. Birds were classified as juvenile or adult based on feather molt pattern (Cramp, 1988).

2.2 | Molecular methods

DNA was extracted using a salt extraction protocol (Richardson, Jury, Blaakmeer, Komdeur, & Burke, 2001). Sexing PCRs (Griffiths, Double, Orr, & Dawson, 1998) determined the sex of the bird and confirmed that DNA extractions were successful.

2.3 | Parasite screening

We used a nested PCR approach that detects *Plasmodium* and *Haemoproteus* to characterize malaria infection status (Waldenström, Bensch, Hasselquist, & Östman, 2004), with multiple positive and negative controls included in each PCR plate. Samples that successfully amplified at least twice were classified as infected. The strain of *Plasmodium* was determined by Sanger sequencing; *Haemoproteus* was not detected. We focused on *Plasmodium* as the most widespread and abundant haemosporidian found in Berthelot's pipits (Illera et al., 2008; Spurgin et al., 2012). In addition, the vector species of *Plasmodium* (mosquitoes; Culicidae) and *Leucocytozoon* (blackflies; Simuliidae) have different ecological niches (Harrigan et al., 2014; Imura et al., 2012), thus combining the two may confound results.

2.4 | Sanger sequencing and SNP genotyping

A previous TLR4 study genotyped 23–30 individuals from all 13 Berthelot's pipit populations, including Tenerife and Porto Santo (González-Quevedo et al., 2015). The primers *PauTLR4F*, *PauTLR4R* (Grueber, Wallis, King, & Jamieson, 2012) were used to amplify a 660-bp region located within the leucine-rich repeat domain of TLR4. Five SNPs were found: TLR4_1 (905 bp, nonsynonymous), TLR4_2 (970 bp, nonsynonymous), TLR4_3 (990 bp, synonymous), TLR4_4 (992 bp, nonsynonymous, triallelic), and TLR4_5 (1,010 bp, nonsynonymous). Base-pair positions are stated according to the zebra finch TLR4 protein coding region, GenBank accession FJ695612. Further Sanger sequencing of TLR4 was performed on all samples collected from Porto Santo in 2016 (*n* = 129). TLR4 SNPs were called through visual inspection of chromatograms in FinchTV (https://digitalworldbiology.com/FinchTV). We used LGC Genomics' proprietary KASP™ genotyping technology (https://www.lgcgroup.com)

to genotype all additional samples from Tenerife and Porto Santo (*n* = 577) at each TLR4 SNP, except for TLR4_5 which was excluded due to a very low minor allele frequency of <0.05. Assay design and genotyping were performed by LGC Genomics, Hertfordshire.

A GWAS performed on restriction site-associated DNA sequencing (RAD-seq) data from Berthelot's pipits by Armstrong et al. (2018) detected two SNPs (5239s1, Chr10:12048280; 7259s1, Chr20:6483195; SNP positions on zebra finch genome v3.2.4, Warren et al., 2010) that showed significant associations with *Plasmodium* strain LK6 (Ortego, Calabuig, Cordero, & Aparicio, 2007) infection. All birds were genotyped with KASP™ assays at these two SNPs.

2.5 | Genetic analysis

Genotypes AT, CT, and TT at the triallelic TLR4_4 SNP were coded as missing data (Tenerife *n* = 12; Porto Santo *n* = 1), to treat this SNP as biallelic. We used DnaSP v6 (Librado & Rozas, 2009) to phase the TLR4 SNPs into haplotypes. To aid phasing, we included TLR4 sequences from all Berthelot's pipit populations, and each phased TLR4 haplotype previously detected in Berthelot's pipits (González-Quevedo et al., 2015). Samples with <90% phasing certainty were excluded from models that included TLR4 haplotypes as predictors. We translated the phased TLR4 sequences originating from Sanger sequencing into protein haplotypes. This gave us the amino acid residues at each of the codons containing a SNP, from which we were able to infer the amino acids at each SNP for samples that were genotyped with KASP™ genotyping.

We used PLINK 1.9 (Chang et al., 2015) to calculate linkage disequilibrium (LD) between each pair of SNPs, and test for deviations from Hardy–Weinberg equilibrium with Hardy–Weinberg exact tests. Where frequencies ≤0.05 were found for SNP minor alleles or TLR4 protein haplotypes, these variants were not included as predictors in genetic analyses.

2.6 | Porto Santo GIS analyses

A previous study demonstrated the importance of environmental and anthropogenic factors in shaping malaria risk in Tenerife (González-Quevedo et al., 2014). Variables were selected based on the potential for influencing the abundance of avian malaria or its mosquito vectors: the minimum temperature of the coldest month (MINTEMP), annual precipitation (PRECIPITATION), altitude (ALTITUDE), aspect (ASPECT), slope (SLOPE), pipit density (DENSITY), vegetation type (VEGTYPE), distance to nearest poultry farm (DISTPOUL), distance to nearest livestock farm (DISTFARM), distance to nearest artificial water reservoir (DISTWATER), and distance to urban site (DIST_URB). We calculated the values of these variables at each sampling location on Porto Santo as outlined below.

All GIS analyses were performed in QGIS v2.18 (QGIS Development Team, 2017). MINTEMP and PRECIPITATION were obtained from WorldClim global climate data v2 (Fick & Hijmans, 2017) at a resolution of 30 arc-seconds (approximately 1 km^2). ALTITUDE

was obtained from the Shuttle Radar Topography Mission (SRTM) 3 Arc-Second Global elevation data (srtm.csi.cgiar.org) at a resolution of approximately 90 m^2. SLOPE and ASPECT were calculated from SRTM data. VEGTYPE was characterized using the CORINE Land Cover inventory (CLC 2012 v.18.5.1; http://land.copernicus.eu/pan-european/corine-land-cover/clc-2012). We combined the land cover classes into six categories: arable, urban-associated, forest, rock-associated, grass, and shrub (Table 1). Values of MINTEMP, PRECIPITATION, ALTITUDE, SLOPE, and ASPECT were calculated by taking the average value of each variable within a 100 m buffer around each sample location. In the case of VEGTYPE, the sample was assigned the category with the largest area within the buffer. ASPECT was classified as one of eight categories: N, NW, W, SW, S, SE, E, and NE.

We calculated DISTWATER with polygons drawn on Google Earth satellite imagery over water sources encountered during sample collection and obtained from the OpenStreetMap data-filtering tool Overpass Turbo (https://overpass-turbo.eu) using the query "natural = water." Water sources were 76–28,190 m^2. The presence of livestock was dependent on visual encounters, as farming census data were not publicly available. The type of livestock was used to differentiate between factors related to livestock farming that might cause aggregations of birds (DISTFARM), and the potential effect of poultry as a reservoir of avian malaria (DISTPOUL). DIST_URB was the distance to the nearest urban-associated area, as classified by VEGTYPE.

DENSITY was calculated as follows. A 1 km^2 grid was overlaid on Porto Santo, and a 1 km radius buffer was drawn around the centroid

TABLE 1 CORINE Land Cover (CLC 2012 v.18.5.1; http://land.copernicus.eu/pan-european/corine-land-cover/clc-2012) classes used to categorize vegetation type on Porto Santo

Vegetation type	CLC classes
Arable	Non-irrigated arable land
	Vineyards
	Complex cultivation patterns
	Land principally occupied by agriculture with significant areas of natural vegetation
Urban-associated	Discontinuous urban fabric
	Port areas
	Airports
	Sport and leisure (resort complex)
Forest	Coniferous forest
Rock-associated	Beaches, dunes, and sand
	Bare rock
	Sparsely vegetated areas
Grass	Natural grassland
	Pastures
	Sport and leisure (golf course)
Shrub	Moors and heathland
	Transitional woodland-shrub

TABLE 2 Single-predictor binomial generalized linear models of the environmental predictors of *Plasmodium* strain LK6 infection in adult Berthelot's pipits on Porto Santo

Variable	Estimate	R^{2a}	*p*-Value	AICc
ALTITUDE	−0.0160	0.129	<.001	100.0
DISTWATER	−0.0013	0.118	<.001	101.2
DIST_URB	−0.0014	0.113	<.001	101.8
VEGTYPE	—	0.157	.170	103.4
MINTEMP	2.5188	0.080	.005	105.4
SLOPE	−0.1569	0.069	.008	106.6
PRECIPITATION	−0.0308	0.041	.035	109.7
DISTPOUL	−0.0006	0.018	.152	112.2
DISTFARM	−0.0005	0.010	.284	113.1
DENSITY	0.0747	0.003	.586	113.9
ASPECT	—	0.094	.271	117.2

Note: Environmental variables are ordered by increasing AICc scores. Parameter estimates are not included for categorical variables.
[a]McFadden pseudo-R^2.

of each grid cell. Any part of the buffer covering the ocean was removed. This was converted into a measure of pipits per km^2 for each grid cell by dividing the number of samples within the buffer by its area. The average density within a 100 m buffer around each sample, weighted by the area of each 1 km^2 grid cell occupied within the sample buffer, was used as the DENSITY measure.

2.7 | Environmental predictors of malaria risk

We modeled the environmental predictors of malaria infection on Porto Santo and used the predicted values from this model as a measure of the malaria risk at each sample location. This allowed us to account for spatial variation in the likelihood of malaria exposure when analyzing the relationship between genotype and infection status. This was calculated using a model selection and then model averaging approach. Model selection tests all combinations of variables as predictors of the response variable and calculates the Akaike information criterion (AIC; Akaike, 1973) as a measure of fit of each model. Model averaging is then applied to a set of models

with the lowest AIC (and therefore the highest likelihood) to calculate weighted averages of parameter estimates and the relative importance of model predictors (Burnham & Anderson, 2002). We report AICc, a modification of AIC that is recommended for small sample sizes (Hurvich & Tsai, 1989).

We used variance inflation factors (VIFs) calculated using the R package car (Fox & Weisberg, 2011) to test for collinearity between environmental variables, using a threshold of >3 to indicate unacceptably high collinearity (Zuur, Ieno, & Elphick, 2010). As we had categorical variables, we used generalized VIFs (GVIFs; Fox & Monette, 1992) transformed with ($GVIF^{1/2df}$)^2 (*df* = degrees of freedom), to calculate a value equivalent to a standard VIF (Fox & Weisberg, 2011). When including all variables, every variable except ASPECT had VIFs > 3 (collinearity). We sequentially removed variables with VIFs > 3 that had the highest AICc scores from single-predictor binomial generalized linear models (GLMs) of each variable against malaria infection status (Table 2). Variables were removed and VIFs recalculated until all variables had VIFs < 3. The remaining variables were VEGTYPE, ALTITUDE, DISTWATER, DENSITY, ASPECT, and DISTPOUL.

Interactions between environmental variables may have biologically meaningful influences on malaria risk. For each biologically relevant pair of variables (Table 3), we tested whether the inclusion of an interaction term improved the fit of a binomial GLM with the two variables as main effects and malaria infection as the response. The interaction DENSITY*DISTWATER gave the largest improvement in AICc (main effects only, AICc = 103.3; main effects and interaction, AICc = 90.8) and was therefore included in model selection (Table 3).

2.8 | Model selection and model averaging

Fitting all combinations of the six environmental variables and one interaction term (see above) as predictors of malaria infection using binomial GLMs, we performed model selection and model averaging following Grueber, Nakagawa, Laws, and Jamieson (2011) using the R package MuMIn (Bartoń, 2018) to obtain the best-supported models for explaining occurrence of malaria infection. Prior to analysis, we used the R package arm (Gelman & Su, 2018) to standardize the input variables to a mean of zero and a standard deviation of 0.5

	Main effect AICc	Interaction AICc	ΔAICc	Retained
DENSITY*DISTWATER	103.3	90.8	−12.5	✓
DISTPOUL*DISTWATER	103.1	99.6	−3.5	
SLOPE*ALTITUDE	100.2	97.1	−3.1	
ALTITUDE*DISTWATER	92.3	91.2	−1.1	
ALTITUDE*DISTPOUL	101.2	101.3	0.1	
DENSITY*DISTPOUL	114.3	114.7	0.4	
ASPECT*SLOPE	116.4	118.3	1.9	
VEGTYPE*DENSITY	105.6	107.7	2.2	

TABLE 3 The effect of including biologically relevant interaction terms between environmental variables for predicting *Plasmodium* LK6 infection in adult Berthelot's pipits on Porto Santo

Note: Binomial generalized linear models were performed with each pair of variables as main effects only, or including an interaction term. Where the addition of an interaction resulted in a change in AICc (ΔAICc) < −7, that interaction term was included in model selection.

to enable meaningful comparisons of parameter estimates (Gelman, 2008; Grueber et al., 2011). The model selection process calculated ΔAICc, the difference in AICc between each model and the "best" model (the model with the lowest AICc), and the Akaike weight, which quantifies the likelihood of each model having the best explanatory power within a set of models (Burnham & Anderson, 2002). Using the R package DescTools (Signorell, 2018), we calculated the McFadden-adjusted pseudo-R^2 (the likelihood of a logistic regression model relative to an intercept-only model, adjusted to account for the number of predictors in the model; McFadden, 1974). Values of McFadden R^2 between 0.2 and 0.4 represent a strong fit, equivalent to a linear regression R^2 of 0.7–0.9 (Louviere, Hensher, & Swait, 2000).

A threshold of ΔAICc ≤ 7 is recommended to retain models that have sufficient support, without dismissing models which still provide some explanatory power (Burnham, Anderson, & Huyvaert, 2011). We applied model averaging over this set of models to calculate weighted averages of parameter estimates and the relative importance of each predictor (the sum of Akaike weights for models which include that predictor). We used the zero method of model averaging to avoid biasing results toward predictors with low explanatory power (Burnham & Anderson, 2002; Lukacs et al., 2007).

2.9 | Spatial autocorrelation

We tested for spatial autocorrelation in model residuals as this may lead to spurious associations between predictor and response variables (Dormann et al., 2007; Lennon, 2000). We created Moran's *I* correlograms at distance class intervals of 750 m and 1,000 m using the R package ncf (Bjornstad, 2018), with 1,000 permutations to test the significance of Moran's *I* at each interval. After correcting for multiple testing using the Holm correction (Holm, 1979; Legendre & Legendre, 2012), there was no evidence of spatial autocorrelation in the model residuals (all adjusted *p* values > .05). Correcting for spatial autocorrelation was therefore not required for the estimation of malaria risk in Porto Santo.

2.10 | Malaria risk scores

We used the predicted values of the best model identified by model selection as a malaria risk score between 0 and 1 for each sample location. This represented the probability of an individual at that location being infected with malaria, as a result of the environmental conditions. An earlier study determined that malaria infection in Tenerife was best explained by DISTPOUL, DISTWATER, MINTEMP, and DISTWATER*MINTEMP (González-Quevedo et al., 2014). As significant spatial autocorrelation was present in Tenerife, an autocovariate term was included in all model combinations during model selection, to account for autocorrelation up to 1,000 m (see González-Quevedo et al., 2014). Hence, the predicted values from an autologistic model containing these predictors, interaction, and autocovariate were used as our estimate of malaria

risk for Tenerife. Malaria risk was logit-transformed prior to use in models.

2.11 | Genetic associations with malaria infection

Genetic variation was classified in three ways: (a) SNP genotype, encoded as 1 for heterozygotes and 0 or 2 for each of the homozygotes; (b) presence (1) or absence (0) of each TLR4 protein haplotype; (c) SNP heterozygosity, with heterozygotes encoded as 1 and homozygotes as 0. Each model described below was performed three times, using each of the three classes of genetic variants as model predictors.

We first tested for associations between genetic variation and malaria infection status. We ran separate models for each island as the genetic divergence between Berthelot's pipits on Tenerife and Porto Santo, along with environmental differences in malaria risk patterns between the islands, might otherwise obscure genetic associations. We tested for genetic associations with malaria infection across all years using binomial generalized linear mixed models (GLMMs) with sampling year as a random effect, within Tenerife (2011) or Porto Santo (2016) using binomial GLMs. Testing across all years gave the advantage of a larger sample size; however, temporal fluctuations in selection pressures could interfere with genetic associations with malaria, so single-year models were also performed.

We tested the genetic variables as predictors of malaria risk in general linear models (LMs) for Tenerife (2011) and Porto Santo (2016). As malaria risk was derived from spatially varying environmental predictors, genetic variation alone was unlikely to account for all spatial autocorrelation in malaria risk. We therefore included distance-based Moran's eigenvector maps (dbMEMs) as spatial predictors of malaria risk. dbMEMs are used to identify gradients of spatial variation (spatial structure) in a response variable, across multiple potential scales from broad to fine, calculated by eigenvector decomposition of distance matrices based on the spatial coordinates of samples (Borcard & Legendre, 2002; Dray, Legendre, & Peres-Neto, 2006). Hence, the use of dbMEMs as predictors of malaria risk accounts for spatially autocorrelated variation in malaria prevalence that would otherwise be explained by environmental conditions, transmission dynamics, and/or unmeasured genetic gradients. We calculated dbMEMs for each island using the R packages adespatial (Dray et al., 2018) and vegan (Oksanen et al., 2018), retaining dbMEMs with positive eigenvalues, representing positively autocorrelated spatial variation. dbMEMs were ranked by descending R^2 values in single-predictor LMs of malaria risk and sequentially added into each model of the genetic associations with malaria risk outlined above, until additional dbMEMs no longer improved AICc. With each iteration, we checked whether spatial autocorrelation in model residuals had been controlled for, to find the minimum required number of dbMEMs. The inclusion of dbMEMs resulted in VIFs < 3, indicating that any collinearity between dbMEMs and genetic variants was acceptably low. We performed hierarchical partitioning using the "lmg" method in the

R package relaimpo (Grömping, 2006) to calculate the proportion of variance in malaria risk explained by genetic variants.

3 | RESULTS

3.1 | Sequencing

Malaria was detected in 126 out of 190 individuals (66.3%) from Porto Santo and 189 out of 590 individuals (32.0%) from Tenerife (Table 4). All infected samples had one of four strains of *Plasmodium*, with no evidence of multiple infection. Between 2006 and 2010, only LK6 (Ortego et al., 2007) was detected (Illera et al., 2008; Spurgin et al., 2012). In samples from Tenerife in 2011, where *Plasmodium* was found in 148 of 388 individuals (38.1%), the majority of infections were LK6 (139 samples; 93.9%). Strains LK5 (Ortego et al., 2007) and KYS9 (Inci et al., 2012) were present in seven (4.7%) and two (1.4%) individuals, respectively (González-Quevedo et al., 2014). Out of 129 samples collected from Porto Santo in 2016, 97 (75.2%) were infected. Of these, 91 (93.8%) had LK6, four (4.1%) had LK5, and a previously undocumented strain, PS1530, was found in two (2.1%) individuals. To avoid potential confounding factors arising from selection against different *Plasmodium* strains, we have focused here on the predominant LK6 strain.

Excluding sampling years where no juveniles were caught, the prevalence of malaria was significantly higher in adults than in juveniles, both in Porto Santo (test of equal proportions $\chi^2 = 25.6$, $p < 0.001$) and in Tenerife ($\chi^2 = 3.1$, $p = 0.039$). As juveniles were present in much lower numbers than adults (Table 4), we removed juveniles from further analysis. Final sample sizes are shown in Table 4.

Sanger sequencing of TLR4 in the 2016 Porto Santo samples only detected SNPs which have been previously documented in Berthelot's pipits (González-Quevedo et al., 2015). SNP allele frequencies differed between the two islands, with different minor alleles found at TLR4_3 and TLR4_4 (Figure 2a). Phasing of the TLR4 SNPs produced five nucleotide haplotypes (Table 5), all of which had been previously detected (González-Quevedo et al., 2015). As amino acid substitutions could potentially alter TLR4 function (Schröder & Schumann, 2005), we classified TLR4 variation into four protein haplotypes (denoted with the prefix "TLR4_P"; Table 5 and Figure 2b). TLR4_P2 was translated from two haplotypes differing at the synonymous SNP TLR4_3. The TLR4_P3 and TLR4_P4 haplotypes were absent from Tenerife, and TLR4_P3 was at low frequency (<0.05) in Porto Santo.

We tested for deviations from Hardy–Weinberg equilibrium at each SNP for each combination of island and year, and in each island across all years. SNPs 5239s1 (Tenerife 2011 and all years), 7259s1 (Tenerife 2006 and 2009), and TLR4_1 (Porto Santo 2009) showed deviations from Hardy–Weinberg equilibrium at $p < .05$; however, following Holm correction for multiple comparisons, all adjusted p values were > 0.05. LD between SNPs TLR4_3 and TLR4_4 (excluding the low-frequency T allele) was high (Porto Santo $R^2 = 0.99$;

TABLE 4 *Plasmodium* prevalence in Berthelot's pipits from Porto Santo and Tenerife across sampling years

Island	Year	n	Infected	n Adults	Infected adults	n Juveniles	Infected juveniles	LK6	LK5	KYS9	PS1530	n Final
Porto Santo		190	0.66	142	0.77	48	0.35	0.95	0.03	0.00	0.02	136
	2006	31	0.65	10	0.90	21	0.52	1.00	0.00	0.00	0.00	10
	2009	30	0.30	16	0.38	14	0.21	1.00	0.00	0.00	0.00	16
	2016	129	0.75	116	0.81	13	0.23	0.94	0.04	0.00	0.02	110
Tenerife		590	0.32	528	0.34	62	0.16	0.95	0.04	0.01	0.00	519
	2006	51	0.08	51	0.08	0	0.00	1.00	0.00	0.00	0.00	51
	2009	56	0.20	25	0.16	31	0.23	1.00	0.00	0.00	0.00	25
	2010	96	0.27	65	0.35	31	0.10	1.00	0.00	0.00	0.00	65
	2011	387	0.38	387	0.38	0	0.00	0.94	0.05	0.01	0.00	378

Note: Sample sizes are given for all birds (*n*), adults (*n* adults), and juveniles (*n* juveniles). The proportion of birds infected with any strain of *Plasmodium* is listed for all birds, adults, and juveniles. The different strains of *Plasmodium* found in Berthelot's pipits (LK6, LK5, KYS9, PS1530) are displayed as proportions of the infected samples. *n* final shows the sample size of the final dataset used in analyses of malaria risk and genetic associations after filtering out juveniles and samples infected with strains other than LK6.
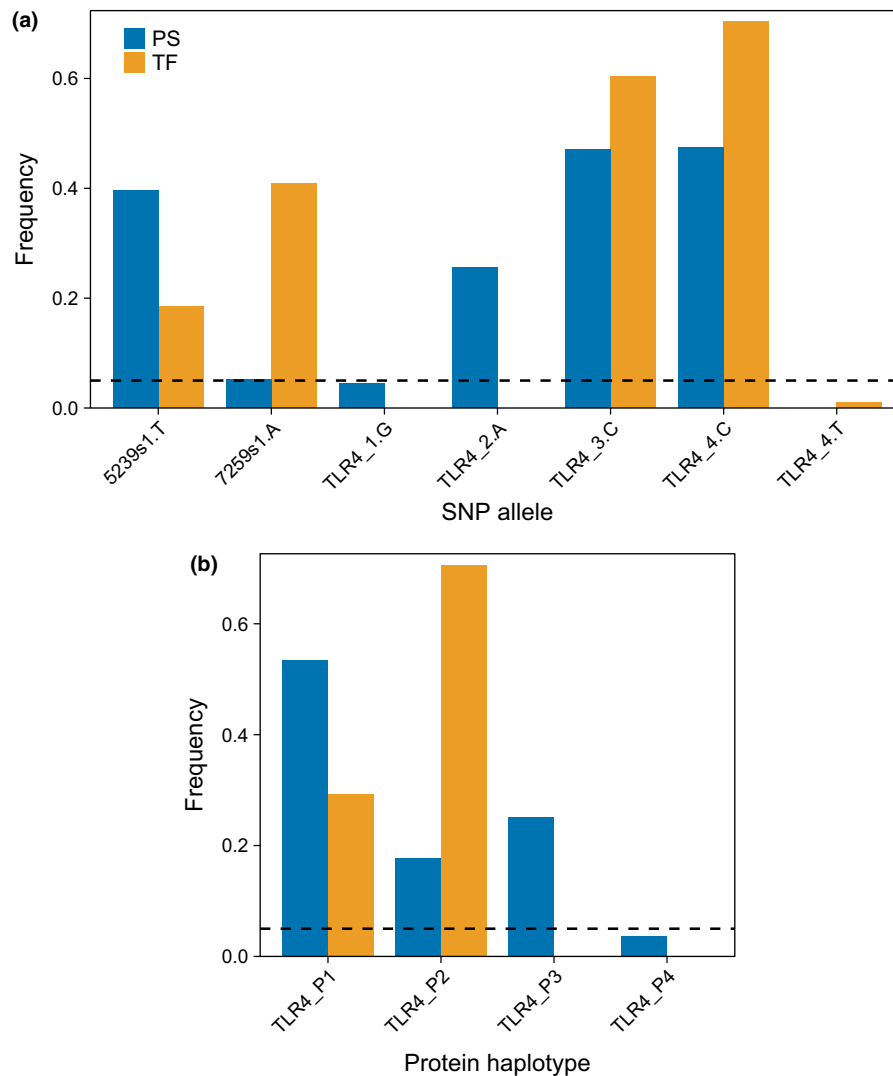
FIGURE 2 Genetic variant frequencies in adult Berthelot's pipits on Porto Santo (PS; blue) and Tenerife (TF; orange). (a) Allele frequencies per SNP. In each instance, the allele which is the minor allele in Porto Santo is represented. In addition, the low-frequency T allele of the triallelic SNP TLR4_4 is also shown. (b) Frequencies of TLR4 protein alleles. In both plots, the black dashed line indicates a frequency threshold of 0.05; variants below this threshold were excluded from models of genetic associations with malaria risk and infection status

TABLE 5 TLR4 nucleotide and protein haplotypes in Berthelot's pipits on Tenerife and Porto Santo

Haplotype	Sequence	Protein haplotype	Amino acid sequence
1	AGTA	TLR4_P1	DGPK
2	AGCC	TLR4_P2	DGPT
3	AGTC	TLR4_P2	DGPT
4	AACC	TLR4_P3	DDPT
5	GGCC	TLR4_P4	GGPT

Tenerife $R^2 = 0.62$). We found moderate LD in Porto Santo between SNPs TLR4_2 and TLR4_3 ($R^2 = 0.39$) and between TLR4_2 and TLR4_4 ($R^2 = 0.37$). All other combinations of SNPs had low levels of LD ($R^2 < 0.1$).

3.2 | Malaria risk models

Model selection of the environmental predictors of malaria infection found 17 models with ΔAICc ≤ 7 relative to the "best" model, which contained VEGTYPE, ALTITUDE, DISTWATER, DENSITY, and DISTWATER*DENSITY (Table 6). ALTITUDE and DISTWATER were negatively associated with malaria infection, whereas DENSITY was positively associated (Figure 3). A *post hoc* Tukey test of VEGTYPE as a predictor of malaria infection (using the R package multcomp; Hothorn, Bretz, & Westfall, 2008) found that rock-associated habitat had a significantly negative effect on malaria infection relative to arable ($p = 0.015$) and grass habitats ($p = 0.039$; Figure 3). These four predictors, and the interaction term DISTWATER*DENSITY, had relative importances of 0.63–0.99 across the top model set (Table 7). We used the predicted values from the best model as our estimate for malaria risk for

TABLE 6 Parameter estimates of predictors included in model selection of the environmental predictors of *Plasmodium* LK6 infection in adult Berthelot's pipits on Porto Santo

Model	Intercept	VEGTYPE	ALTITUDE	DENSITY	DISTPOUL	DISTWATER	SLOPE	DENSITY* DISTWATER	AICc	ΔAICc	Akaike weight	Adjusted R^{2a}
1	2.62	+	-1.91	0.63		-1.54		-5.33	84.9	0.0	0.29	0.25
2	2.61	+	-1.75	0.40	-0.82	-1.39		-5.85	86.4	1.5	0.14	0.24
3	1.97		-1.32	1.01		-2.47		-4.17	86.4	1.5	0.13	0.22
4	2.63	+	-2.00	0.74		-1.56	0.31	-5.28	87.2	2.3	0.09	0.23
5	1.94		-1.23	0.79	-0.69	-2.18		-4.74	87.8	2.9	0.07	0.21
6	1.94		-1.08	0.83		-2.38	-0.52	-4.25	88.1	3.2	0.06	0.21
7	2.62	+	-1.96	0.59	-1.00	-1.34	0.70	-5.89	88.5	3.6	0.05	0.22
8	1.77			0.60		-2.36	-1.21	-4.79	89.0	4.1	0.04	0.20
9	1.93		-1.06	0.68	-0.62	-2.14	-0.42	-4.71	89.7	4.8	0.03	0.20
10	1.76			0.47	-0.65	-2.09	-1.05	-5.22	90.4	5.5	0.02	0.19
11	1.98	+		0.65		-1.81		-5.45	90.5	5.6	0.02	0.19
12	1.74			0.94		-2.43		-4.62	90.8	5.9	0.01	0.18
13	2.01	+		0.35	-1.09	-1.43		-6.20	91.0	6.1	0.01	0.19
14	1.72			0.66	-0.99	-2.04		-5.45	91.1	6.2	0.01	0.18
15	2.32	+	-2.45						91.3	6.4	0.01	0.18
16	2.06	+		0.36		-1.93	-0.98	-5.61	91.4	6.6	0.01	0.19
17	2.32	+	-2.12			-1.35			91.5	6.6	0.01	0.18

Note: Models within a threshold of ΔAICc ≤ 7 relative to the best model are shown.
[a]McFadden-adjusted pseudo-R^2.

FIGURE 3 Environmental predictors of *Plasmodium* strain LK6 risk in adult Berthelot's pipits on Porto Santo. (a) The proportion of samples infected with LK6 per category of VEGTYPE. * indicates significant difference between categories in post hoc Tukey tests at *p* < 0.05. Plots b–d show logistic regression models of the effect of (b) altitude, (c) distance to water sources, and (d) pipit density on the probability of LK6 infection. Histograms indicate the frequency of uninfected (lower) and infected (upper) individuals for each (b) 50 m altitude class, (c) 200 m distance class, and (d) 0.5 pipits/km^2 class. Shaded area represents 95% confidence intervals
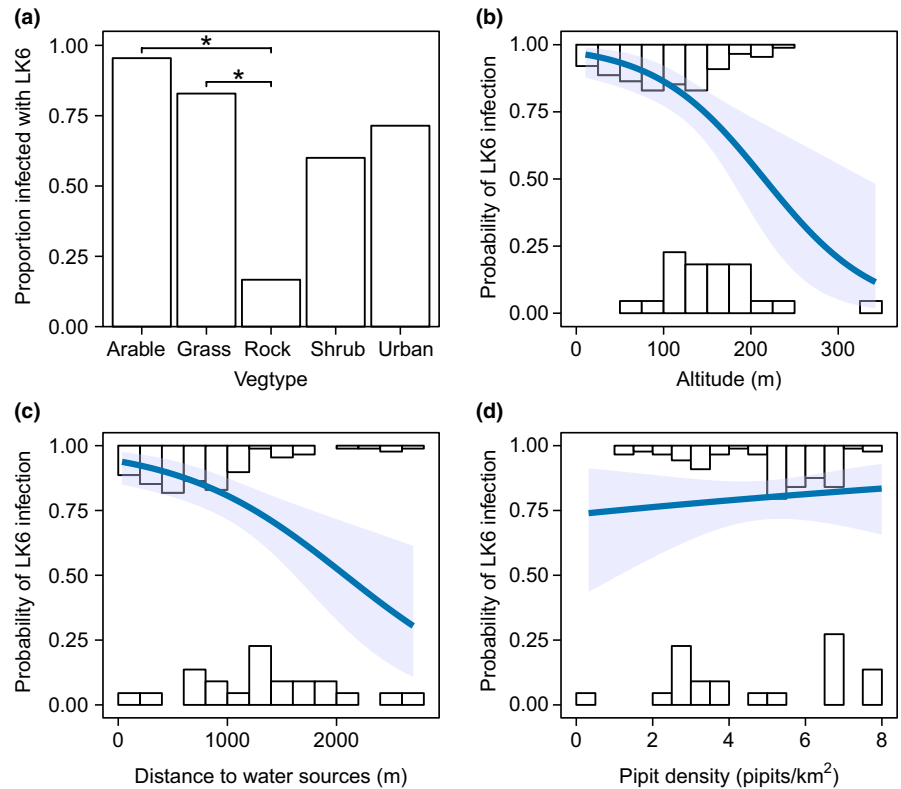
each sample location (Figure 4). DISTPOUL and SLOPE had low relative importance in model averaging (0.33 and 0.29, respectively), and ASPECT did not feature within the top model set. The malaria risk model for Porto Santo had a McFadden-adjusted R^2 of 0.25 (values of 0.2–0.4 are equivalent to a linear regression R^2 of 0.7–0.9; Louviere et al., 2000), whereas the adjusted R^2 for the Tenerife malaria risk model was 0.10 (González-Quevedo et al.,

2014). Distributions of malaria risk differed markedly between the islands, with higher malaria risk in Porto Santo (Figure 5).

## 3.3 | Genetic associations with malaria infection

As we found high levels of LD between TLR4 SNPs, we calculated VIFs for models of genetic associations with malaria infection and

**TABLE 7** Environmental predictors of malaria risk in adult Berthelot's pipits on Porto Santo

	Best model	Model averaging	Relative importance
(Intercept)	2.62 (1.62, 4.04)	2.32 (1.10, 3.55)	
DISTWATER	−1.54 (−3.64, 0.67)	−1.79 (−3.83, 0.24)	0.99
DENSITY	0.63 (−0.59, 1.91)	0.66 (−0.64, 1.97)	0.98
DENSITY*DISTWATER	−5.33 (−9.01, −2.19)	−5.02 (−8.72, −1.32)	0.98
ALTITUDE	−1.91 (−3.85, −0.54)	−1.47 (−3.33, 0.39)	0.87
VEGTYPE	+	+	0.63
arable	−0.56 (−2.97, 2.58)	−0.31 (−2.42, 1.80)	
rock	−3.77 (−7.55, −1.10)	−2.42 (−6.91, 2.06)	
shrub	−3.37 (−7.28, −0.00)	−1.95 (−6.00, 2.11)	
urban	−1.59 (−3.85, 0.80)	−0.94 (−3.26, 1.38)	
DISTPOUL		−0.26 (−1.47, 0.94)	0.33
SLOPE		−0.05 (−1.27, 1.16)	0.29
Adjusted R^{2a}	0.25		

Note: Parameter estimates (and 95% confidence intervals) are presented for the "best" model with lowest AICc as determined by model selection, and for model averaging across 17 models with ΔAICc ≤ 7 relative to the best model. Within VEGTYPE, coefficients of arable, rock, shrub, and urban were calculated relative to grass as the reference category.
[a]McFadden-adjusted pseudo-R^2.
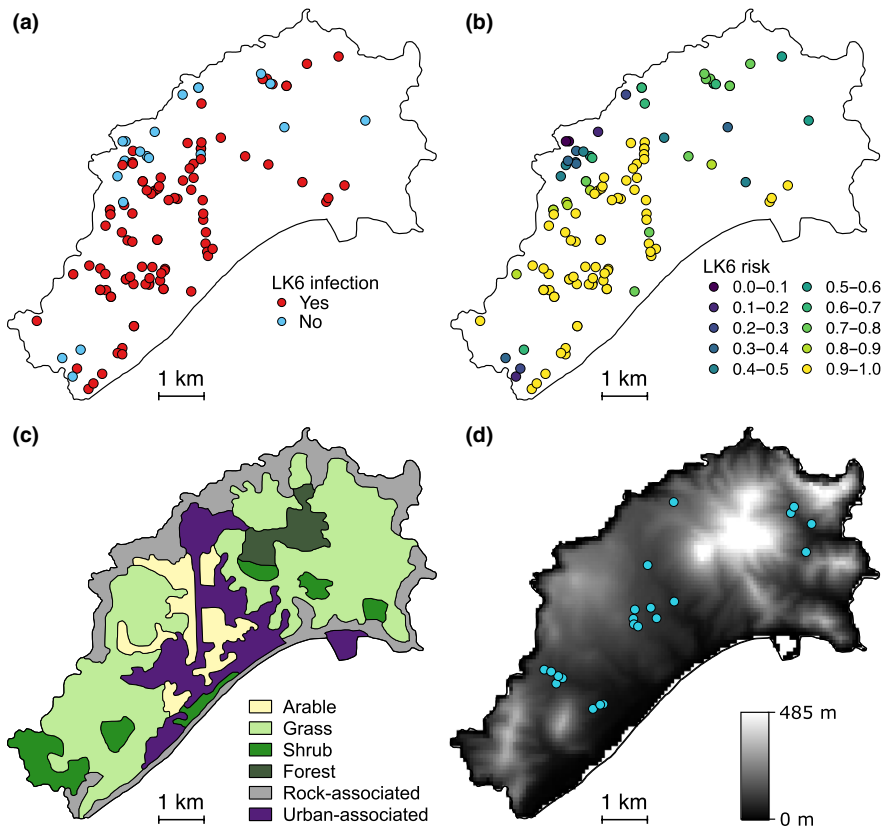
FIGURE 4 Spatial distribution of *Plasmodium* strain LK6 and environmental risk factors in adult Berthelot's pipits on Porto Santo. Points indicate location of (a) individual infection status and (b) infection risk in adult Berthelot's pipits. (c) VEGTYPE categories, adapted from the CORINE Land Cover inventory (CLC 2012 v.18.5.1; http://land.copernicus.eu/pan-european/corine-land-cover/clc-2012). (d) Altitude on Porto Santo, calculated from Shuttle Radar Topography Mission global elevation data (SRTM 90 m; https://srtm.csi.cgiar.org). Overlaid blue points indicate locations of standing water sources
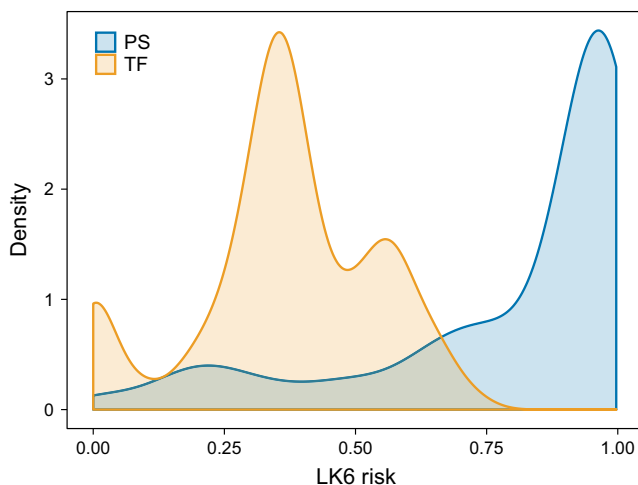
FIGURE 5 The distribution of *Plasmodium* strain LK6 infection risk in adult Berthelot's pipits on Porto Santo (PS; blue) and Tenerife (TF; orange)

risk. All VIF scores were <3 in Tenerife. In Porto Santo, TLR4_3 and TLR4_4 had elevated VIF scores (>50); however, after removing the synonymous TLR4_3 SNP, all VIF scores were <3.

Results from models of genetic variants as predictors of malaria infection are summarized in Table 8. In Porto Santo, across all years, increasing frequency of the T allele in SNP 5239s1 was associated with increased malaria infection (estimate = 0.75, SE = 0.34, p = 0.026; Figure 6a). This effect was no longer significant when looking only at 2016 (p = 0.099), although the direction of the result

remained consistent, with the highest probability of malaria prevalence in TT genotype individuals. This may be a power issue (n = 110 for 2016 vs. n = 136 for all years). No other SNP was significantly associated with malaria infection in 2016 or across all years. There were no associations between malaria infection and SNP heterozygosity or TLR4 protein haplotypes in Porto Santo.

TLR4_P1 presence had a negative effect on malaria infection on Tenerife in 2011 (estimate = −0.46, SE = 0.23, p = 0.041; Figure 6b). We found the same trend across all years, although the association was not significant (p = 0.091). The TLR4_P2 haplotype was not associated with malaria infection in 2011 or across all years. We found no associations with malaria infection for SNP genotypes or SNP heterozygosity in Tenerife.

3.4 | Genetic associations with malaria risk

We tested for associations between genetic variants and malaria risk on Porto Santo (2016) and Tenerife (2011). The results are summarized in Table 9. On Porto Santo, increasing numbers of T alleles at SNP 5239s1 (estimate = 0.69, SE = 0.27, p = 0.011), and A alleles at SNP TLR4_2 (estimate = 1.03, SE = 0.42, p = 0.016), were associated with increased malaria risk. However, the residuals of this model were highly spatially autocorrelated. To control for this, we included seven dbMEMs with high R^2 in single-predictor models of malaria risk (Figure 7), chosen from a set of dbMEMs which gave the lowest AICc in a multipredictor model of malaria risk. After controlling for autocorrelation, SNP 5239s1 was still associated with malaria

TABLE 8 Summary of general linear models of the association between genetic variants and malaria infection status in Berthelot's pipits on Porto Santo (PS) and Tenerife (TF)

Island	Variant type	Variant	All years estimate	Main year estimate
PS	SNP genotype	5239s1	0.75 (0.34)*	0.60 (0.37)
		TLR4_2	0.90 (0.51)	1.10 (0.58)
		TLR4_4	0.15 (0.40)	−0.16 (0.45)
	SNP heterozygosity	5239s1	0.53 (0.45)	0.61 (0.52)
		TLR4_2	0.94 (0.49)	0.95 (0.55)
		TLR4_4	−0.76 (0.47)	−0.98 (0.53)
	TLR4 protein haplotype	TLR4_P1	−1.62 (0.85)	−1.57 (0.94)
		TLR4_P2	−0.53 (0.52)	−0.82 (0.60)
		TLR4_P3	0.40 (0.51)	0.24 (0.59)
TF	SNP genotype	5239s1	0.08 (0.17)	0.18 (0.18)
		7259s1	−0.02 (0.14)	−0.03 (0.15)
		TLR4_3	−0.10 (0.22)	0.09 (0.26)
		TLR4_4	0.31 (0.25)	0.19 (0.29)
	SNP heterozygosity	5239s1	0.09 (0.21)	0.25 (0.24)
		7259s1	−0.29 (0.20)	−0.30 (0.22)
		TLR4_3	0.20 (0.25)	0.00 (0.29)
		TLR4_4	−0.41 (0.26)	−0.40 (0.29)
	TLR4 protein haplotype	TLR4_P1	−0.34 (0.20)	−0.46 (0.23)*
		TLR4_P2	−0.07 (0.38)	−0.20 (0.43)

Note: Parameter estimates (with standard error in brackets) for each genetic variant were taken from multipredictor models with genetic variants coded as SNP heterozygosity, SNP genotype, or TLR4 protein haplotype presence/absence. Models were performed in each island across all sampling years ("All years estimate"), or in the main sampling year with largest sample size (Porto Santo = 2016; Tenerife = 2011; "Main year estimate"). Asterisks next to parameter estimates denote significance of the predictor (*$p < 0.05$).

risk (estimate = 0.38, SE = 0.17, p = 0.030; Figure 8) but TLR4_2 was not (p = 0.423). We did not find an association between TLR4_4 and malaria risk, either before or after controlling for autocorrelation. Hierarchical partitioning of the above models (Table 10) found that SNP 5239s1 explained 5.2% of the variance in malaria risk before controlling for autocorrelation, and 3.3% of the variance after the addition of dbMEMs. Despite having a non-significant association with malaria risk in the model containing dbMEMs, TLR4_2 explained a greater proportion of the variance in malaria risk compared to 5239s1, both before (7.4%) and after (4.3%) controlling for autocorrelation.

Before taking autocorrelation into account, there were no significant associations between SNP heterozygosity and malaria risk on Porto Santo; however, after including seven dbMEMs to remove spatial autocorrelation in model residuals, heterozygosity at SNP 5239s1 was strongly associated with increased malaria risk (estimate = 0.74, SE = 0.25, p = 0.004). Heterozygosity at other SNPs was not associated with malaria risk.

Prior to controlling for autocorrelation, the presence of protein haplotype TLR4_P1 was associated with reduced malaria risk on Porto Santo (estimate = −1.32, SE = 0.62, p = 0.037). The same seven dbMEMs were used to control for spatial autocorrelation, after which there was no longer an effect of TLR4_P1 presence (p = 0.318). No other TLR4 protein haplotypes showed associations with malaria risk.

To investigate the loss of significance of the associations between TLR4_P1 or TLR4_2 and malaria risk on Porto Santo, we ran a binomial GLM (TLR4_P1) and LM (TLR4_2) with the seven dbMEMs as predictors for the two genetic variants. dbMEM1 was significantly associated with both TLR4_P1 (estimate = −1.18, SE = 0.39, p = 0.003) and TLR4_2 (estimate = 0.16, SE = 0.06, p = 0.005).

We did not find any significant associations between SNP genotype, SNP heterozygosity, or TLR4 protein haplotypes and malaria risk on Tenerife. We were unable to remove spatial autocorrelation in model residuals through the addition of dbMEMs as model predictors.

4 | DISCUSSION

We used previously identified candidate SNPs linked to malaria infection across populations (from a GWAS analysis performed on RAD-seq SNPs; Armstrong et al., 2018) and TLR4 SNPs (González-Quevedo et al., 2015) to investigate the relationship between potentially adaptive genetic variation and avian malaria within two island populations of Berthelot's pipits. In addition to testing for associations with infection status, we calculated the malaria risk at each sampling location, predicted by modeling fine-scale environmental
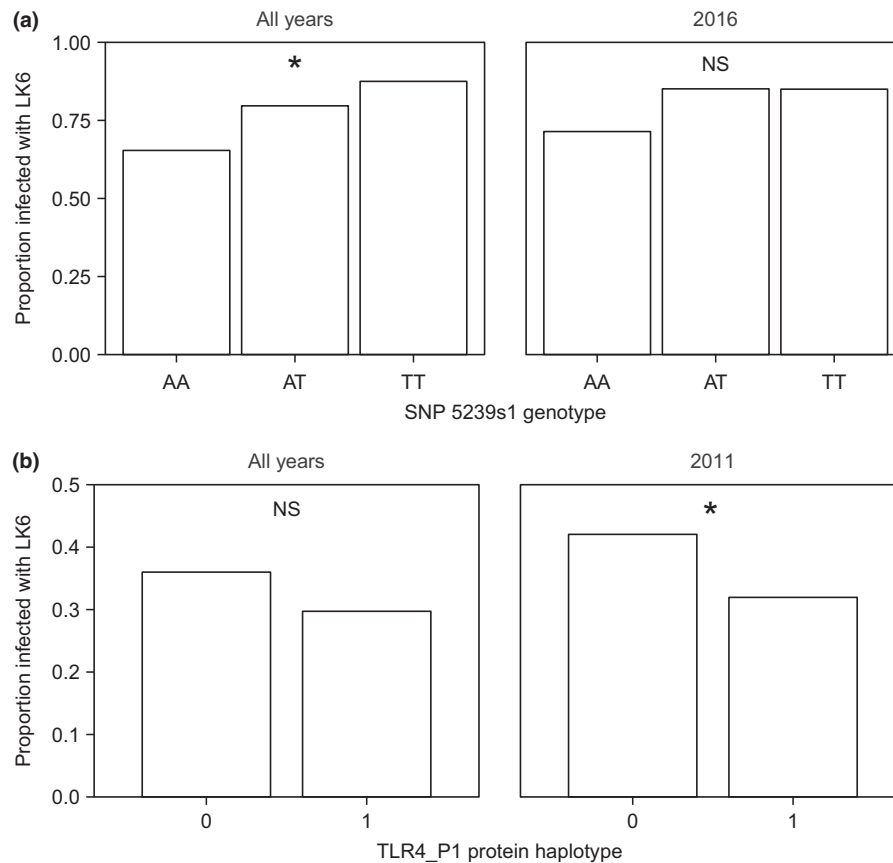
FIGURE 6 The association between *Plasmodium* strain LK6 infection status and genetic variants in adult Berthelot's pipits. (a) SNP 5239s1 genotype association with LK6 in Porto Santo. (b) Protein haplotype TLR4_P1 association with LK6 in Tenerife. 0 = absent, 1 = present. *Denotes significance at *p* < 0.05; NS = not significant

drivers of malaria infection. We found associations between malaria infection status and SNP 5239s1 in Porto Santo, and TLR4 protein haplotype 1 in Tenerife. Furthermore, the SNPs 5239s1 and TLR4_2 showed associations with malaria risk in Porto Santo, but not in Tenerife, where malaria risk was lower.

4.1 | Genetic associations with malaria

We have previously used RAD-seq SNPs to detect genetic variants that were associated with LK6 infection in Berthelot's pipits in the Canary Islands (Armstrong et al., 2018). The strongest association was found for SNP 5239s1, ca. 2,000 bp from interleukin-16, a pro-inflammatory cytokine that moderates the expression of other cytokines associated with malaria infection (Kern, Hemmer, Damme, Gruss, & Dietrich, 1989; Lyke et al., 2004; Mathy et al., 2000). In the present study, SNP 5239s1 was a predictor of malaria on Porto Santo, with the lowest infection and risk found in samples with the AA genotype. Remarkably, this was the opposite relationship to that found in the Canary Islands (Armstrong et al., 2018), where increased incidence of the T allele was associated with reduced infection. This may be indicative of pathogen-mediated balancing selection, which can arise from heterozygote advantage (Doherty & Zinkernagel, 1975), rare-allele advantage (Slade & McCallum, 1992; Takahata & Nei, 1990), and local adaptation to fluctuating

pathogen selection pressures (Hill et al., 1991). When controlling for spatial autocorrelation, we found an association between SNP 5239s1 heterozygosity and malaria risk on Porto Santo, although contrary to the heterozygote advantage model, heterozygotes were associated with greater malaria risk than homozygotes (an effect which was largely driven by the decline in risk found with AA genotypes). Berthelot's pipit populations on the Madeiran and Canary Islands archipelagos have been isolated from each other for at least 8,500 years (Spurgin et al., 2014). Different populations may therefore be undergoing independent coevolutionary cycles with the same malaria strain, with alternative alleles conferring an advantage between divergent populations (Bonneaud, Pérez-Tris, Federici, Chastel, & Sorci, 2006). Alternatively, undetected genetic and phenotypic differences within the LK6 strain could potentially drive local adaptation between the archipelagos, with different alleles favored in different populations (Alcaide, Edwards, Negro, Serrano, & Tella, 2008; Loiseau et al., 2009). We used a single genetic marker, the mitochondrial cytochrome b locus, to classify the malaria strain. Several genes on the *Plasmodium* genome with relevance to infection success have shown greater genetic variation than at cytochrome b (Jarvi, Farias, & Atkinson, 2008; Lauron et al., 2014). It is possible that Berthelot's pipits on separate archipelagos could be adapting to different malaria strains within LK6, although this remains to be tested.

TABLE 9 Summary of linear models of the association between genetic variants and malaria risk in Berthelot's pipits on Porto Santo (PS) and Tenerife (TF)

Island	Variant type	Variant	Estimate	dbMEM estimate
PS	SNP genotype	5239s1	0.69 (0.27)*	0.38 (0.17)*
		TLR4_2	1.03 (0.42)*	0.22 (0.27)
		TLR4_4	0.21 (0.36)	0.09 (0.23)
	SNP heterozygosity	5239s1	0.79 (0.41)	0.74 (0.25)**
		TLR4_2	0.77 (0.42)	0.39 (0.26)
		TLR4_4	−0.59 (0.42)	−0.36 (0.26)
	TLR4 protein haplotype	TLR4_P1	−1.32 (0.62)*	−0.41 (0.41)
		TLR4_P2	−0.36 (0.51)	0.02 (0.33)
		TLR4_P3	0.60 (0.47)	0.14 (0.30)
TF	SNP genotype	5239s1	0.26 (0.73)	
		7259s1	−0.23 (0.60)	
		TLR4_3	−0.03 (1.04)	
		TLR4_4	0.40 (1.13)	
	SNP heterozygosity	5239s1	0.80 (0.96)	
		7259s1	−1.08 (0.86)	
		TLR4_3	0.48 (1.13)	
		TLR4_4	−0.59 (1.14)	
	TLR4 protein haplotype	TLR4_P1	−0.45 (0.89)	
		TLR4_P2	0.22 (1.68)	

Note: Parameter estimates (with standard error in brackets) for each genetic variant were taken from multipredictor models with genetic variants coded as SNP heterozygosity, SNP genotype, or TLR4 protein haplotype presence/absence. On Porto Santo, models were performed with just the genetic variants, or with the inclusion of dbMEMs to control for autocorrelation in model residuals ("dbMEM estimate"). dbMEMs were unable to account for autocorrelation in Tenerife models. Asterisks next to parameter estimates denote significance of the predictor (*p <0 .05; **p < 0.01).

We did not find evidence of associations between SNP 5239s1 and malaria infection or risk on Tenerife, despite this population being included in the previous GWAS (Armstrong et al., 2018). SNPs that are related to individual-level variation in parasite burden do not necessarily show the same associations at the landscape scale (Wenzel, Douglas, James, Redpath, & Piertney, 2016). It is possible that with the comparatively low malaria risk in Tenerife, gene flow is overriding landscape-scale associations between SNP 5239s1 and malaria risk (Forester, Jones, Joost, Landguth, & Lasky, 2016; Lenormand, 2002). The previous GWAS result could have been driven by other populations such as Lanzarote and Fuerteventura, where malaria infection rates were higher (Illera et al., 2008; Spurgin et al., 2012).

Polymorphisms in immune genes can alter the effectiveness of their proteins for detecting and responding to pathogens (Lazarus et al., 2002; Sommer, 2005). The TLR4 SNPs sequenced here are situated within the ligand-binding region, which plays a key role in TLR pathogen recognition (Werling, Jann, Offord, Glass, & Coffey, 2009). Evidence of positive selection in birds or mammals has been detected at each of the codons identified as polymorphic in Berthelot's pipits (Areal, Abrantes, & Esteves, 2011; Králová et al., 2018; Wlasiuk & Nachman, 2010), suggesting that these sites may be important for the evolution of pathogen recognition. On Tenerife, the presence of the TLR4 protein haplotype TLR4_P1 was associated with decreased malaria infection prevalence in 2011, but not across all sampling

years. In earlier years, approximately half of the samples were collected from the high-altitude (>2,000 m above sea level) plateau of El Teide. Malaria has not been found in Berthelot's pipits in this location (González-Quevedo et al., 2014; Illera et al., 2008; Spurgin et al., 2012), although a survey of passerine communities on Tenerife found malaria at low frequency in high-altitude habitats (Padilla et al., 2017). The relationship seen in 2011 between TLR4_P1 and infection may be masked in other sampling years by the increase in uninfected individuals from areas of low malaria abundance. We did not find a relationship between TLR4_P1 presence and malaria risk, potentially due to the explanatory power of the Tenerife malaria risk model (McFadden-adjusted pseudo-R^2 = 0.10). On Porto Santo, both TLR4_P1 and the SNP TLR4_2 were associated with malaria risk, although these relationships were no longer significant after including dbMEMs to remove autocorrelation. Both of these genetic variants showed significant associations with dbMEM1, which itself explained 22% of the variance in malaria risk, making it difficult to disentangle the real effects of these variants from any spurious associations arising from residual autocorrelation.

4.2 | Environmental predictors of malaria risk

We modeled the environmental predictors of malaria distributions in Porto Santo to understand fine-scale spatial differences in malaria
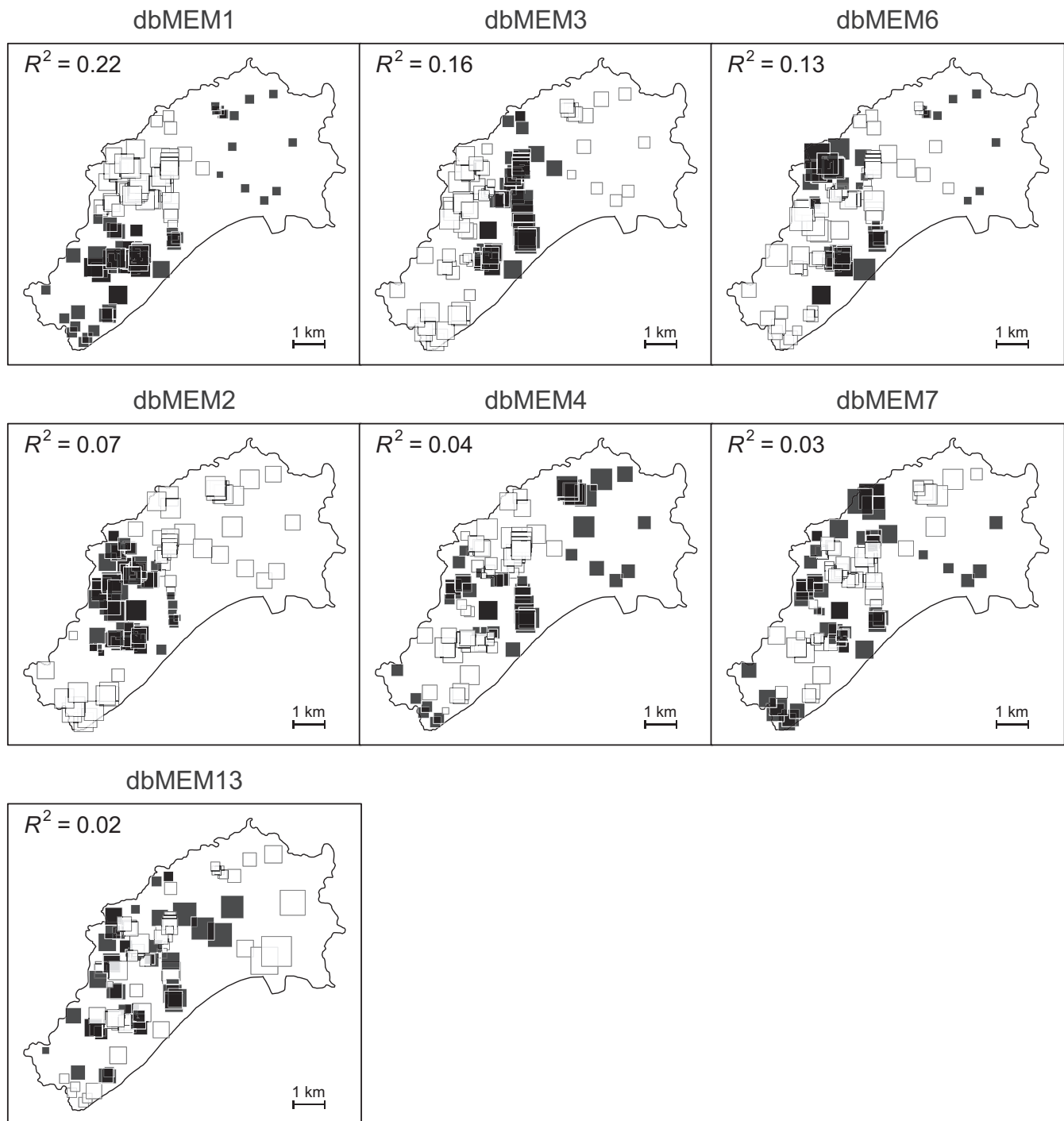
FIGURE 7 Distance-based Moran's eigenvector maps (dbMEMs) showing strongest associations with *Plasmodium* strain LK6 infection risk in adult Berthelot's pipits on Porto Santo. R^2 values for each dbMEM were calculated from single-predictor LMs of each dbMEM against logit-transformed LK6 risk. Positive eigenvector scores are indicated by black squares, and negative scores are white. The size of the square indicates the magnitude of the score

risk. Higher altitudes were associated with decreased probability of malaria infection on Porto Santo, whereas on Tenerife, temperature was a predictor of malaria. Collinearity between altitude, temperature, and precipitation was found on both islands, with the same climatic processes likely influencing malaria distributions (González-Quevedo et al., 2014). This is perhaps not surprising as malaria vector distributions are constrained by thermal requirements, with

decreased malaria prevalence often reported at high altitudes (Eggert et al., 2008; Niebuhr, Poulin, & Tompkins, 2016) and low temperatures (Blanford et al., 2013; Craig, Le Sueur, & Snow, 1999; Loiseau et al., 2013).

Distance to water sources was an important predictor of malaria distributions in both Porto Santo and Tenerife (González-Quevedo et al., 2014). Due to the aquatic larval development of mosquitoes,
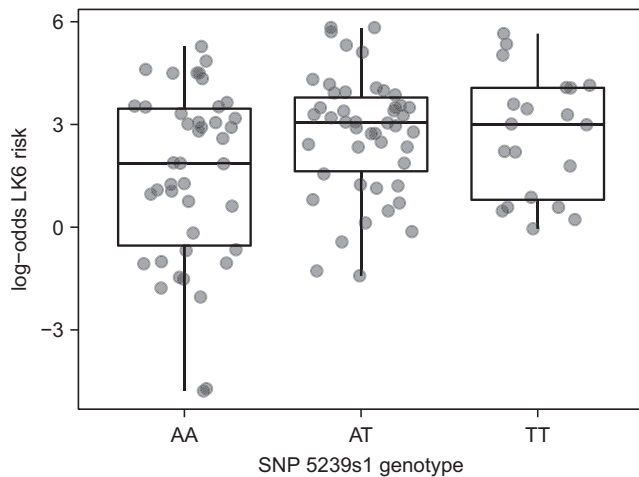
FIGURE 8 The association between *Plasmodium* strain LK6 infection risk and SNP 5239s1 in adult Berthelot's pipits on Porto Santo. LK6 risk was logit-transformed prior to model fitting; a log-odds risk score of 0 is equivalent to an infection probability of .5

TABLE 10 Hierarchical partitioning of variance for predictors of malaria risk in Berthelot's pipits on Porto Santo

Predictor	R^2	R^2 (with dbMEMs)
5239s1	0.052	0.033
TLR4_2	0.074	0.043
TLR4_4	0.039	0.027
dbMEMs		0.589

Note: R^2 values were calculated from a general linear model with genotypes at SNPs 5239s1, TLR4_2, and TLR4_4 as predictors of malaria risk, with and without the inclusion of dbMEMs as predictors to control for autocorrelation.

higher vector abundance and malaria are found in proximity to water (Ferraguti et al., 2018; Ganser et al., 2016; Illera et al., 2017). In the present study, distance to urban areas was removed prior to model selection due to a positive collinearity with distance to water sources. Therefore, we cannot rule out the importance of additional sources of standing water that may be associated with urban environments. Other studies have found links between urbanization and increased malaria and/or vector abundance (Alemu, Tsegaye, Golassa, & Abebe, 2011; Li et al., 2014), although this appears to vary between vector species, with some favoring more natural habitats (Ferraguti et al., 2016). Pipit density was positively associated with malaria risk on Porto Santo, although the model-averaged parameter estimate was relatively small. There was, however, a strong negative interaction between distance to water and pipit density on this island, likely due to aggregations of mosquitoes and hosts around water sources, which may increase disease transmission rates (Begon et al., 2002; Greer, Briggs, & Collins, 2008; Le Menach, McKenzie, Flahault, & Smith, 2005; Raghwani et al., 2011).

Vegetation type was associated with malaria prevalence on Porto Santo. The highest abundance of malaria was found in arable and grassland habitats, with lower malaria in rock-associated habitats. However, this result should be interpreted with caution due to small sample sizes, as only six pipits were caught on rock-associated habitats. While not an important predictor of malaria infection in Tenerife (González-Quevedo et al., 2014), differences in malaria and vector abundances between vegetation types have been found elsewhere (Clark, Wells, Dimitrov, & Clegg, 2016; Ferreira Junior et al., 2017; Rubio-Palis & Zimmerman, 1997).

Contrary to findings from Tenerife (González-Quevedo et al., 2014), distance to poultry was not an important predictor of malaria prevalence on Porto Santo. This may be because the effects of poultry farms as disease reservoirs (either due to the poultry themselves or due to aggregations of wild birds around them), that are driving increased malaria abundance on Tenerife (González-Quevedo et al.,

2014), do not have an effect at the small scale of poultry farming witnessed on Porto Santo.

By testing for associations with malaria infection and risk at TLR4 and novel malaria-associated SNPs in divergent populations, we have revealed contrasting patterns of malaria risk and potential local adaptation, potentially due to different patterns of coevolution between the two populations. In addition, we found genetic associations with environmentally driven fine-scale spatial variation in malaria risk at the landscape scale within Porto Santo. A lack of genetic associations with malaria risk in Tenerife may indicate the importance of spatial scales for assessing local adaptation across landscapes, where fine-scale associations may be obscured over larger areas. Understanding the processes of local adaptation and the environmental drivers of infectious disease will be of additional importance for conservation efforts, as future climatic fluctuations alter the prevalence of disease.

ACKNOWLEDGMENTS

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

DSR, RGD, and LGS designed the research, obtained funding, and supervised the project. DSR, LGS, CG-Q, and CA collected samples.

ORCID

Claire Armstrong https://orcid.org/0000-0003-3874-4243

Catalina González-Quevedo https://orcid.org/0000-0003-0256-6721

Molly Dunne https://orcid.org/0000-0001-6267-536X

Lewis G. Spurgin https://orcid.org/0000-0002-0874-9281

David S. Richardson https://orcid.org/0000-0001-7226-9074

REFERENCES

Akaike, H. (1973). Information theory as an extension of the maximum likelihood principle. In B. N. Petrov, & F. Csaki (Eds.), *Second international symposium on information theory* (pp. 267–281). Budapest, Hungary: Akademiai Kiado.

Alcaide, M., Edwards, S. V., Negro, J. J., Serrano, D., & Tella, J. L. (2008). Extensive polymorphism and geographical variation at a positively selected MHC class II B gene of the lesser kestrel (*Falco naumanni*). *Molecular Ecology*, *17*(11), 2652–2665. https://doi.org/10.1111/j.1365-294X.2008.03791.x

Alemu, A., Tsegaye, W., Golassa, L., & Abebe, G. (2011). Urban malaria and associated risk factors in Jimma town, south-west Ethiopia. *Malaria Journal*, *10*(1), 173. https://doi.org/10.1186/1475-2875-10-173

Areal, H., Abrantes, J., & Esteves, P. J. (2011). Signatures of positive selection in Toll-like receptor (TLR) genes in mammals. *BMC Evolutionary Biology*, *11*(1), 368. https://doi.org/10.1186/1471-2148-11-368

Armstrong, C., Richardson, D. S., Hipperson, H., Horsburgh, G. J., Küpper, C., Percival-Alwyn, L., … Spurgin, L. G. (2018). Genomic associations with bill length and disease reveal drift and selection across island bird populations. *Evolution Letters*, *2*, 22–36. https://doi.org/10.1002/evl3.38

Atkinson, C. T., Saili, K. S., Utzurrum, R. B., & Jarvi, S. I. (2013). Experimental evidence for evolved tolerance to avian malaria in a wild population of low elevation Hawai'i 'Amakihi (*Hemignathus virens*). *EcoHealth*, *10*(4), 366–375. https://doi.org/10.1007/s10393-013-0899-2

Bartoń, K. (2018). *MuMIn: Multi-Model Inference. R package version 1.40.4*. Retrieved from https://CRAN.R-project.org/package=MuMIn

Begon, M., Bennett, M., Bowers, R. G., French, N. P., Hazel, S. M., & Turner, J. (2002). A clarification of transmission terms in host-microparasite models: Numbers, densities and areas. *Epidemiology and Infection*, *129*(1), 147–153. https://doi.org/10.1017/S0950268802007148

Bernatchez, L., & Landry, C. (2003). MHC studies in nonmodel vertebrates: What have we learned about natural selection in 15 years? *Journal of Evolutionary Biology*, *16*(3), 363–377. https://doi.org/10.1046/j.1420-9101.2003.00531.x

Bjornstad, O. N. (2018). *ncf: Spatial Covariance Functions. R package version 1.2-4*. Retrieved from https://cran.r-project.org/package=ncf

Blanford, J. I., Blanford, S., Crane, R. G., Mann, M. E., Paaijmans, K. P., Schreiber, K. V., & Thomas, M. B. (2013). Implications of temperature variation for malaria parasite development across Africa. *Scientific Reports*, *3*(1), 1300. https://doi.org/10.1038/srep01300

Bockelmann, A. C., Reusch, B. H., Bijlsma, R., & Bakker, J. P. (2003). Habitat differentiation vs. isolation-by-distance: The genetic population structure of *Elymus athericus* in European salt marshes. *Molecular Ecology*, *12*(2), 505–515. https://doi.org/10.1046/j.1365-294X.2003.01706.x

Bonneaud, C., Pérez-Tris, J., Federici, P., Chastel, O., & Sorci, G. (2006). Major histocompatibility alleles associated with local resistance to malaria in a passerine. *Evolution*, *60*(2), 383–389. https://doi.org/10.1111/j.0014-3820.2006.tb01114.x

Borcard, D., & Legendre, P. (2002). All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modelling*, *153*(1–2), 51–68. https://doi.org/10.1016/S0304-3800(01)00501-4

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic-approach* (2nd ed.). Berlin, Germany: Springer.

Burnham, K. P., Anderson, D. R., & Huyvaert, K. P. (2011). AIC model selection and multimodel inference in behavioral ecology: Some background, observations, and comparisons. *Behavioral Ecology and Sociobiology*, *65*(1), 23–35. https://doi.org/10.1007/s00265-010-1029-6

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*(1), 7. https://doi.org/10.1186/s13742-015-0047-8

Clark, N. J., Clegg, S. M., & Lima, M. R. (2014). A review of global diversity in avian haemosporidians (*Plasmodium* and *Haemoproteus*: Haemosporida): New insights from molecular data. *International Journal for Parasitology*, *44*(5), 329–338. https://doi.org/10.1016/j.ijpara.2014.01.004

Clark, N. J., Wells, K., Dimitrov, D., & Clegg, S. M. (2016). Co-infections and environmental conditions drive the distributions of blood parasites in wild birds. *Journal of Animal Ecology*, *85*(6), 1461–1470. https://doi.org/10.1111/1365-2656.12578

Craig, M. H., Le Sueur, D., & Snow, R. W. (1999). A climate-based distribution model of malaria transmission in sub-Saharan Africa. *Parasitology Today*, *15*(3), 105–111. https://doi.org/10.1016/S0169-4758(99)01396-4

Cramp, S. (1988). *The birds of the Western Palearctic* (vol. *5*). Oxford, UK: Oxford University Press.

Doherty, P. C., & Zinkernagel, R. M. (1975). Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature*, *256*, 50–52. https://doi.org/10.1038/256050a0

Dormann, C. F., McPherson, J. M., Araújo, M. B., Bivand, R., Bolliger, J., & Carl, G., … Wilson, R. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography*, *30*(5), 609–628. https://doi.org/10.1111/j.2007.0906-7590.05171.x

Dray, S., Bauman, D., Blanchet, G., Borcard, D., Clappe, S., Guenard, G., & Wagner, H. H. (2018). *adespatial: Multivariate Multiscale Spatial Analysis. R package version 0.2-0*. Retrieved from https://CRAN.R-project.org/package=adespatial

Dray, S., Legendre, P., & Peres-Neto, P. R. (2006). Spatial modelling: A comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*, *196*(3–4), 483–493. https://doi.org/10.1016/j.ecolmodel.2006.02.015

Eggert, L. S., Terwilliger, L. A., Woodworth, B. L., Hart, P. J., Palmer, D., & Fleischer, R. C. (2008). Genetic structure along an elevational gradient in Hawaiian honeycreepers reveals contrasting evolutionary

responses to avian malaria. *BMC Evolutionary Biology*, 8, 315. https://doi.org/10.1186/1471-2148-8-315

Fallon, S. M., Bermingham, E., & Ricklefs, R. E. (2005). Host specialization and geographic localization of avian malaria parasites: A regional analysis in the Lesser Antilles. *The American Naturalist*, 165(4), 466–480. https://doi.org/10.1086/428430

Fellay, J., Shianna, K. V., Ge, D., Colombo, S., Ledergerber, B., Weale, M., … Goldstein, D. B. (2007). A whole-genome association study of major determinants for host control of HIV-1. *Science*, 317(5840), 944–947. https://doi.org/10.1126/science.1143767

Ferraguti, M., Martínez-de la Puente, J., Bensch, S., Roiz, D., Ruiz, S., Viana, D. S., … Figuerola, J. (2018). Ecological determinants of avian malaria infections: An integrative analysis at landscape, mosquito and vertebrate community levels. *Journal of Animal Ecology*, 87(3), 727–740. https://doi.org/10.1111/1365-2656.12805

Ferraguti, M., Martínez-De La Puente, J., Roiz, D., Ruiz, S., Soriguer, R., & Figuerola, J. (2016). Effects of landscape anthropization on mosquito community composition and abundance. *Scientific Reports*, 6(1), 29002. https://doi.org/10.1038/srep29002

Ferreira Junior, F. C., Rodrigues, R. A., Ellis, V. A., Leite, L. O., Borges, M. A., & Braga, E. M. (2017). Habitat modification and seasonality influence avian haemosporidian parasite distributions in southeastern Brazil. *PLoS ONE*, 12(6), e0178791. https://doi.org/10.1371/journal.pone.0178791

Ferrer-Admetlla, A., Bosch, E., Sikora, M., Marquès-Bonet, T., Ramírez-Soriano, A., Muntasell, A., … Casals, F. (2008). Balancing selection is the main force shaping the evolution of innate immunity genes. *The Journal of Immunology*, 181(2), 1315–1322. https://doi.org/10.4049/jimmunol.181.2.1315

Ferwerda, B., McCall, M. B. B., Alonso, S., Giamarellos-Bourboulis, E. J., Mouktaroudi, M., Izagirre, N., … Netea, M. G. (2007). TLR4 polymorphisms, infectious diseases, and evolutionary pressure during migration of modern humans. *Proceedings of the National Academy of Sciences of the United States of America*, 104(42), 16645–16650. https://doi.org/10.1073/pnas.0704828104

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302–4315. https://doi.org/10.1002/joc.5086

Fisher, C. A., Bhattarai, E. K., Osterstock, J. B., Dowd, S. E., Seabury, P. M., Vikram, M., … Seabury, C. M. (2011). Evolution of the bovine TLR gene family and member associations with *Mycobacterium avium* subspecies *paratuberculosis* infection. *PLoS ONE*, 6(11), e27744. https://doi.org/10.1371/journal.pone.0027744

Forester, B. R., Jones, M. R., Joost, S., Landguth, E. L., & Lasky, J. R. (2016). Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Molecular Ecology*, 25(1), 104–120. https://doi.org/10.1111/mec.13476

Fox, J., & Monette, G. (1992). Generalized collinearity diagnostics. *Journal of the American Statistical Association*, 87(417), 178–183. https://doi.org/10.1080/01621459.1992.10475190

Fox, J., & Weisberg, S. (2011). *An R companion to applied regression* (2nd ed.). Thousand Oaks, CA: Sage.

Fu, W.-X., Liu, Y., Lu, X., Niu, X.-Y., Ding, X.-D., Liu, J.-F., & Zhang, Q. (2012). A genome-wide association study identifies two novel promising candidate genes affecting *Escherichia coli* F4ab/F4ac susceptibility in swine. *PLoS ONE*, 7(3), e32127. https://doi.org/10.1371/journal.pone.0032127

Fumagalli, M., Sironi, M., Pozzoli, U., Ferrer-Admetlla, A., Pattini, L., & Nielsen, R. (2011). Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLoS Genetics*, 7(11), e1002355. https://doi.org/10.1371/journal.pgen.1002355

Ganser, C., Gregory, A. J., Mcnew, L. B., Hunt, L. A., Sandercock, B. K., & Wisely, S. M. (2016). Fine-scale distribution modeling of avian malaria vectors in north-central Kansas. *Journal of Vector Ecology*, 41(1), 114–122. https://doi.org/10.1111/jvec.12202

Garroway, C. J., Radersma, R., Sepil, I., Santure, A. W., De Cauwer, I., Slate, J., & Sheldon, B. C. (2013). Fine-scale genetic structure in a wild bird population: The role of limited dispersal and environmentally based selection as causal factors. *Evolution*, 67(12), 3488–3500. https://doi.org/10.1111/evo.12121

Gavan, M. K., Oliver, M. K., Douglas, A., & Piertney, S. B. (2015). Gene dynamics of toll-like receptor 4 through a population bottleneck in an insular population of water voles (*Arvicola amphibius*). *Conservation Genetics*, 16, 1181–1193. https://doi.org/10.1007/s10592-015-0731-3

Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine*, 27(15), 2865–2873. https://doi.org/10.1002/sim.3107

Gelman, A., & Su, Y.-S. (2018). *arm: Data Analysis Using Regression and Multilevel/Hierarchical Models. R package version 1.10-1*. Retrieved from https://CRAN.R-project.org/package=arm

González-Quevedo, C., Davies, R. G., Phillips, K. P., Spurgin, L. G., & Richardson, D. S. (2016). Landscape-scale variation in an anthropogenic factor shapes immune gene variation within a wild population. *Molecular Ecology*, 25(17), 4234–4246. https://doi.org/10.1111/mec.13759

González-Quevedo, C., Davies, R. G., & Richardson, D. S. (2014). Predictors of malaria infection in a wild bird population: Landscape-level analyses reveal climatic and anthropogenic factors. *Journal of Animal Ecology*, 83, 1091–1102. https://doi.org/10.1111/1365-2656.12214

González-Quevedo, C., Spurgin, L. G., Illera, J. C., & Richardson, D. S. (2015). Drift, not selection, shapes toll-like receptor variation among oceanic island populations. *Molecular Ecology*, 24(23), 5852–5863. https://doi.org/10.1111/mec.13437

Greer, A. L., Briggs, C. J., & Collins, J. P. (2008). Testing a key assumption of host-pathogen theory: Density and disease transmission. *Oikos*, 117(11), 1667–1673. https://doi.org/10.1111/j.1600-0706.2008.16783.x

Griffiths, R., Double, M. C., Orr, K., & Dawson, R. J. G. (1998). A DNA test to sex most birds. *Molecular Ecology*, 7(8), 1071–1075. https://doi.org/10.1046/j.1365-294x.1998.00389.x

Grömping, U. (2006). Relative importance for linear regression in R: The package relaimpo. *Journal of Statistical Software*, 17, 1–27.

Grueber, C. E., Nakagawa, S., Laws, R. J., & Jamieson, I. G. (2011). Multimodel inference in ecology and evolution: Challenges and solutions. *Journal of Evolutionary Biology*, 24(4), 699–711. https://doi.org/10.1111/j.1420-9101.2010.02210.x

Grueber, C. E., Wallis, G. P., King, T. M., & Jamieson, I. G. (2012). Variation at innate immunity toll-like receptor genes in a bottlenecked population of a New Zealand robin. *PLoS ONE*, 7(9), e45011. https://doi.org/10.1371/journal.pone.0045011

Guggisberg, A. M., Sayler, K. A., Wisely, S. M., & Odom John, A. R. (2018). Natural history of *Plasmodium odocoilei* malaria infection in farmed white-tailed deer. *mSphere*, 3(2), e00067-18. https://doi.org/10.1128/mSphere.00067-18

Harrigan, R. J., Sedano, R., Chasar, A. C., Chaves, J. A., Nguyen, J. T., Whitaker, A., & Smith, T. B. (2014). New host and lineage diversity of avian haemosporidia in the northern Andes. *Evolutionary Applications*, 7(7), 799–811. https://doi.org/10.1111/eva.12176

He, D., Tao, S., Guo, S., Li, M., Wu, J., Huang, H., … Wang, Y. (2015). Interaction of TLR-IFN and HLA polymorphisms on susceptibility of chronic HBV infection in Southwest Han Chinese. *Liver International*, 35(8), 1941–1949. https://doi.org/10.1111/liv.12756

Hereford, J. (2009). A quantitative survey of local adaptation and fitness trade-offs. *The American Naturalist*, 173(5), 579–588. https://doi.org/10.1086/597611

Hill, A. V. S., Allsopp, C. E. M., Kwiatkowski, D., Anstey, N. M., Twumasi, P., Rowe, P. A., … Greenwood, B. M. (1991). Common west African

HLA antigens are associated with protection from severe malaria. *Nature*, *352*(6336), 595–600. https://doi.org/10.1038/352595a0

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(6), 65–70. https://doi.org/10.2307/4615733

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*(3), 346–363. https://doi.org/10.1002/bimj.200810425

Hurvich, C. M., & Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika*, *76*(2), 297–307. https://doi.org/10.1093/biomet/76.2.297

Illera, J. C., Emerson, B. C., & Richardson, D. S. (2007). Population history of Berthelot's pipit: Colonization, gene flow and morphological divergence in Macaronesia. *Molecular Ecology*, *16*(21), 4599–4612. https://doi.org/10.1111/j.1365-294X.2007.03543.x

Illera, J. C., Emerson, B. C., & Richardson, D. S. (2008). Genetic characterization, distribution and prevalence of avian pox and avian malaria in the Berthelot's pipit (*Anthus berthelotii*) in Macaronesia. *Parasitology Research*, *103*(6), 1435–1443. https://doi.org/10.1007/s00436-008-1153-7

Illera, J. C., Fernández-Álvarez, Á., Hernández-Flores, C. N., & Foronda, P. (2015). Unforeseen biogeographical patterns in a multiple parasite system in Macaronesia. *Journal of Biogeography*, *42*, 1858–1870. https://doi.org/10.1111/jbi.12548

Illera, J. C., López, G., García-Padilla, L., & Moreno, Á. (2017). Factors governing the prevalence and richness of avian haemosporidian communities within and between temperate mountains. *PLoS ONE*, *12*(9), e0184587. https://doi.org/10.1371/journal.pone.0184587

Imura, T., Suzuki, Y., Ejiri, H., Sato, Y., Ishida, K., Sumiyama, D., ... Yukawa, M. (2012). Prevalence of avian haematozoa in wild birds in a high-altitude forest in Japan. *Veterinary Parasitology*, *183*(3–4), 244–248. https://doi.org/10.1016/j.vetpar.2011.07.027

Inci, A., Yildirim, A., Njabo, K. Y., Duzlu, O., Biskin, Z., & Ciloglu, A. (2012). Detection and molecular characterization of avian *Plasmodium* from mosquitoes in central Turkey. *Veterinary Parasitology*, *188*(1–2), 179–184. https://doi.org/10.1016/j.vetpar.2012.02.012

Jarvi, S. I., Farias, M. E. M., & Atkinson, C. T. (2008). Genetic characterization of Hawaiian isolates of *Plasmodium relictum* reveals mixed-genotype infections. *Biology Direct*, *3*(1), 25. https://doi.org/10.1186/1745-6150-3-25

Jepson, A., Banya, W., Sisay-joof, F., Hassan-King, M., Nunes, C., Bennett, S., & Whittle, H. (1997). Quantification of the relative contribution of major histocompatibility complex (MHC) and non-MHC genes to human immune responses to foreign antigens. *Infection and Immunity*, *65*(3), 872–876.

Jones, M. R., Cheviron, Z. A., & Carling, M. D. (2013). Spatial patterns of avian malaria prevalence in *Zonotrichia capensis* on the western slope of the Peruvian Andes. *The Journal of Parasitology*, *99*(5), 903–905. https://doi.org/10.1645/12-147.1

Kern, P., Hemmer, C. J., Damme, J. V., Gruss, H.-J., & Dietrich, M. (1989). Elevated tumor necrosis factor alpha and interleukin-6 serum levels as markers for complicated *Plasmodium falciparum* malaria. *The American Journal of Medicine*, *87*(2), 139–143. https://doi.org/10.1016/S0002-9343(89)80688-6

Knowles, S. C. L., Palinauskas, V., & Sheldon, B. C. (2010). Chronic malaria infections increase family inequalities and reduce parental fitness: Experimental evidence from a wild bird population. *Journal of Evolutionary Biology*, *23*(3), 557–569. https://doi.org/10.1111/j.1420-9101.2009.01920.x

Králová, T., Albrecht, T., Bryja, J., Hořák, D., Johnsen, A., Lifjeld, J. T., ... Vinkler, M. (2018). Signatures of diversifying selection and convergence acting on passerine Toll-like receptor 4 in an evolutionary context. *Molecular Ecology*, *27*(13), 2871–2883. https://doi.org/10.1111/mec.14724

Langin, K. M., Sillett, T. S., Funk, W. C., Morrison, S. A., Desrosiers, M. A., & Ghalambor, C. K. (2015). Islands within an island: Repeated adaptive divergence in a single population. *Evolution*, *69*(3), 653–665. https://doi.org/10.1111/evo.12610

Larson, W. A., Lisi, P. J., Seeb, J. E., Seeb, L. W., & Schindler, D. E. (2016). Major histocompatibility complex diversity is positively associated with stream water temperatures in proximate populations of sockeye salmon. *Journal of Evolutionary Biology*, *29*(9), 1846–1859. https://doi.org/10.1111/jeb.12926

Lauron, E. J., Oakgrove, K. S., Tell, L. A., Biskar, K., Roy, S. W., & Sehgal, R. N. (2014). Transcriptome sequencing and analysis of *Plasmodium gallinaceum* reveals polymorphisms and selection on the apical membrane antigen-1. *Malaria Journal*, *13*(1), 382. https://doi.org/10.1186/1475-2875-13-382

Lazarus, R., Vercelli, D., Palmer, L. J., Klimecki, W. J., Silverman, E. K., Richter, B., ... Kwiatkowski, D. J. (2002). Single nucleotide polymorphisms in innate immunity genes: Abundant variation and potential role in complex human disease. *Immunological Reviews*, *190*, 9–25. https://doi.org/10.1034/j.1600-065X.2002.19002.x

Le Menach, A., McKenzie, F. E., Flahault, A., & Smith, D. L. (2005). The unexpected importance of mosquito oviposition behaviour for malaria: Non-productive larval habitats can be sources for malaria transmission. *Malaria Journal*, *4*(1), 23. https://doi.org/10.1186/1475-2875-4-23

Legendre, P., & Legendre, L. (2012). *Numerical ecology* (3rd ed.). Amsterdam, The Netherlands: Elsevier.

Lennon, J. J. (2000). Red-shifts and red herrings in geographical ecology. *Ecography*, *23*(1), 101–113. https://doi.org/10.1111/j.1600-0587.2000.tb00265.x

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology and Evolution*, *17*(4), 183–189. https://doi.org/10.1016/S0169-5347(02)02497-7

Levene, H. (1953). Genetic equilibrium when more than one ecological niche is available. *The American Naturalist*, *87*(836), 331–333. https://doi.org/10.1086/281792

Li, Y., Kamara, F., Zhou, G., Puthiyakunnon, S., Li, C., Liu, Y., ... Chen, X.-G. (2014). Urbanization increases *Aedes albopictus* larval habitats and accelerates mosquito development and survivorship. *PLoS Neglected Tropical Diseases*, *8*(11), e3301. https://doi.org/10.1371/journal.pntd.0003301

Librado, P., & Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, *25*(11), 1451–1452. https://doi.org/10.1093/bioinformatics/btp187

Loiseau, C., Harrigan, R. J., Bichet, C., Julliard, R., Garnier, S., Lendvai, Á. Z., ... Sorci, G. (2013). Predictions of avian *Plasmodium* expansion under climate change. *Scientific Reports*, *3*(1), 1126. https://doi.org/10.1038/srep01126

Loiseau, C., Richard, M., Garnier, S., Chastel, O., Julliard, R., Zoorob, R., & Sorci, G. (2009). Diversifying selection on MHC class I in the house sparrow (*Passer domesticus*). *Molecular Ecology*, *18*(7), 1331–1340. https://doi.org/10.1111/j.1365-294X.2009.04105.x

Loiseau, C., Zoorob, R., Robert, A., Chastel, O., Julliard, R., & Sorci, G. (2011). *Plasmodium relictum* infection and MHC diversity in the house sparrow (*Passer domesticus*). *Proceedings of the Royal Society B: Biological Sciences*, *278*(1709), 1264–1272. https://doi.org/10.1098/rspb.2010.1968

Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). *Stated choice methods: Analysis and applications* (p. 402). Cambridge, UK: Cambridge University Press. Retrieved from https://books.google.co.uk/books?hl=en&lr=&xml:id=nk8bpTjutPQC&oi=fnd&pg=PR9&dq=Stated+Choice+Methods:+Analysis+and+Applications&ots=WAUcegankb&sig=Q0ESuHgZMPiXi1ehbO5NKAAh6QU#v=onepage&q=mcfadden&f=false

Lukacs, P. M., Thompson, W. L., Kendall, W. L., Gould, W. R., Doherty, P. F., Burnham, K. P., & Anderson, D. R. (2007). Concerns regarding a call for pluralism of information theory and hypothesis testing. *Journal of Applied Ecology*, *44*(2), 456–460. https://doi.org/10.1111/j.1365-2664.2006.01267.x

Lyke, K. E., Burges, R., Cissoko, Y., Sangare, L., Dao, M., Diarra, I., … Sztein, M. (2004). Serum levels of the proinflammatory cytokines Interleukin-1 Beta (IL-1 β), IL-6, IL-8, IL-10, Tumor Necrosis Factor Alpha, and IL-12 (p70) in Malian children with severe *Plasmodium falciparum* malaria and matched uncomplicated malaria or healthy cont. *Infection and Immunity*, 72(10), 5630–5637. https://doi.org/10.1128/IAI.72.10.5630

MacArthur, R. H., & Wilson, E. O. (1967). *The Theory of island biogeography* (p. 224). Princeton, NJ: Princeton University Press.

Martinsen, E. S., Perkins, S. L., & Schall, J. J. (2008). A three-genome phylogeny of malaria parasites (*Plasmodium* and closely related genera): Evolution of life-history traits and host switches. *Molecular Phylogenetics and Evolution*, 47(1), 261–273. https://doi.org/10.1016/j.ympev.2007.11.012

Marzal, A., Bensch, S., Reviriego, M., Balbontin, J., & De Lope, F. (2008). Effects of malaria double infection in birds: One plus one is not two. *Journal of Evolutionary Biology*, 21(4), 979–987. https://doi.org/10.1111/j.1420-9101.2008.01545.x

Mathy, N. L., Scheuer, W., Lanzendörfer, M., Honold, K., Ambrosius, D., Norley, S., & Kurth, R. (2000). Interleukin-16 stimulates the expression and production of pro-inflammatory cytokines by human monocytes. *Immunology*, 100, 63–69. https://doi.org/10.1046/j.1365-2567.2000.00997.x

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics*. New York, NY: Academic Press.

Mockenhaupt, F. P., Cramer, J. P., Hamann, L., Stegemann, M. S., Eckert, J., Oh, N.-R., … Schumann, R. R. (2006). Toll-like receptor (TLR) polymorphisms in African children: Common TLR-4 variants predispose to severe malaria. *Proceedings of the National Academy of Sciences of the United States of America*, 103(1), 177–182. https://doi.org/10.1073/pnas.0506803102

Netea, M. G., Wijmenga, C., & O'Neill, L. A. (2012). Genetic variation in Toll-like receptors and disease susceptibility. *Nature Immunology*, 13(6), 535–542. https://doi.org/10.1038/ni.2284

Niebuhr, C. N., Poulin, R., & Tompkins, D. M. (2016). Is avian malaria playing a role in native bird declines in New Zealand? Testing hypotheses along an elevational gradient. *PLoS ONE*, 11(11), e0165918. https://doi.org/10.1371/journal.pone.0165918

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., & Helene, W. (2018). *vegan: Community Ecology Package. R package version 2.5-2*. Retrieved from https://CRAN.R-project.org/package=vegan

Olsson-Pons, S., Clark, N. J., Ishtiaq, F., & Clegg, S. M. (2015). Differences in host species relationships and biogeographic influences produce contrasting patterns of prevalence, community composition and genetic structure in two genera of avian malaria parasites in southern Melanesia. *Journal of Animal Ecology*, 84(4), 985–998. https://doi.org/10.1111/1365-2656.12354

Ortego, J., Calabuig, G., Cordero, P. J., & Aparicio, J. M. (2007). Genetic characterization of avian malaria (Protozoa) in the endangered lesser kestrel, *Falco naumanni*. *Parasitology Research*, 101(4), 1153–1156. https://doi.org/10.1007/s00436-007-0575-y

Padilla, D. P., Illera, J. C., González-Quevedo, C., Villalba, M., & Richardson, D. S. (2017). Factors affecting the distribution of haemosporidian parasites within an oceanic island. *International Journal for Parasitology*, 47(4), 225–235. https://doi.org/10.1016/j.ijpara.2016.11.008

Pérez-Rodríguez, A., Ramírez, Á., Richardson, D. S., & Pérez-Tris, J. (2013). Evolution of parasite island syndromes without long-term host population isolation: Parasite dynamics in Macaronesian blackcaps *Sylvia atricapilla*. *Global Ecology and Biogeography*, 22(12), 1272–1281. https://doi.org/10.1111/geb.12084

Piel, F. B., Patil, A. P., Howes, R. E., Nyangiri, O. A., Gething, P. W., Williams, T. N., … Hay, S. I. (2010). Global distribution of the sickle cell gene and geographical confirmation of the malaria hypothesis. *Nature Communications*, 1(8), 104. https://doi.org/10.1038/ncomms1104

QGIS Development Team (2017). *QGIS geographic information system*. Open Source Geospatial Foundation Project.

Raghwani, J., Rambaut, A., Holmes, E. C., Hang, V. T., Hien, T. T., Farrar, J., … Simmons, C. P. (2011). Endemic dengue associated with the co-circulation of multiple viral lineages and localized density-dependent transmission. *PLoS Pathogens*, 7(6), e1002064. https://doi.org/10.1371/journal.ppat.1002064

Ravenhall, M., Campino, S., Sepúlveda, N., Manjurano, A., Nadjm, B., Mtove, G., … Clark, T. G. (2018). Novel genetic polymorphisms associated with severe malaria and under selective pressure in North-eastern Tanzania. *PLoS Genetics*, 14(1), e1007172. https://doi.org/10.1371/journal.pgen.1007172

Richardson, D. S., Jury, F. L., Blaakmeer, K., Komdeur, J., & Burke, T. (2001). Parentage assignment and extra-group paternity in a cooperative breeder: The Seychelles warbler (*Acrocephalus sechellensis*). *Molecular Ecology*, 10(9), 2263–2273. https://doi.org/10.1046/j.0962-1083.2001.01355.x

Richardson, J. L., Urban, M. C., Bolnick, D. I., & Skelly, D. K. (2014). Microgeographic adaptation and the spatial scale of evolution. *Trends in Ecology and Evolution*, 29(3), 165–176. https://doi.org/10.1016/j.tree.2014.01.002

Rubio-Palis, Y., & Zimmerman, R. H. (1997). Ecoregional classification of malaria vectors in the neotropics. *Journal of Medical Entomology*, 34(5), 499–510. https://doi.org/10.1093/jmedent/34.5.499

Schmidt, P. S., Bertness, M. D., & Rand, D. M. (2000). Environmental heterogeneity and balancing selection in the acorn barnacle *Semibalanus balanoides*. *Proceedings of the Royal Society B: Biological Sciences*, 267(1441), 379–384. https://doi.org/10.1098/rspb.2000.1012

Schröder, N. W., & Schumann, R. R. (2005). Single nucleotide polymorphisms of Toll-like receptors and susceptibility to infectious disease. *Lancet Infectious Diseases*, 5(3), 156–164. https://doi.org/10.1016/S1473-3099(05)01308-3

Signorell, A. (2018). *DescTools: Tools for descriptive statistics. R package version 0.99.24*. Retrieved from https://cran.r-project.org/package=DescTools

Slade, R. W., & McCallum, H. I. (1992). Overdominant vs. frequency-dependent selection at MHC loci. *Genetics*, 132, 861–862. https://doi.org/10.1002/path.1711680114/abstract

Sommer, S. (2005). The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Frontiers in Zoology*, 2, 16. https://doi.org/10.1186/1742-9994-2-16

Spurgin, L. G., Illera, J. C., Jorgensen, T. H., Dawson, D. A., & Richardson, D. S. (2014). Genetic and phenotypic divergence in an island bird: Isolation by distance, by colonization or by adaptation? *Molecular Ecology*, 23(5), 1028–1039. https://doi.org/10.1111/mec.12672

Spurgin, L. G., Illera, J. C., Padilla, D. P., & Richardson, D. S. (2012). Biogeographical patterns and co-occurrence of pathogenic infection across island populations of Berthelot's pipit (*Anthus berthelotii*). *Oecologia*, 168(3), 691–701. https://doi.org/10.1007/s00442-011-2149-z

Spurgin, L. G., & Richardson, D. S. (2010). How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B: Biological Sciences*, 277(1684), 979–988. https://doi.org/10.1098/rspb.2009.2084

Takahata, N., & Nei, M. (1990). Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics*, 124, 967–978.

Thye, T., Vannberg, F. O., Wong, S. H., Owusu-Dabo, E., Osei, I., Gyapong, J., … The Wellcome Trust Case Control Consortium (2010). Genome-wide association analyses identifies a susceptibility locus for tuberculosis on chromosome 18q11.2. *Nature Genetics*, 42(9), 739–741. https://doi.org/10.1038/ng.639

Tschirren, B., Andersson, M., Scherman, K., Westerdahl, H., & Råberg, L. (2011). Contrasting patterns of diversity and population differentiation at the innate immunity gene toll-like receptor 2 (TLR2) in two sympatric rodent species. *Evolution*, 66(3), 720–731. https://doi.org/10.5061/dryad.744v6t51

Turcotte, A., Bélisle, M., Pelletier, F., & Garant, D. (2018). Environmental determinants of haemosporidian parasite prevalence in a declining population of Tree swallows. *Parasitology*, 145(7), 961–970. https://doi.org/10.1017/S0031182017002128

Valente, L., Illera, J. C., Havenstein, K., Pallien, T., Etienne, R. S., & Tiedemann, R. (2017). Equilibrium bird species diversity in Atlantic islands. *Current Biology*, 27(11), 1660–1666.e5. https://doi.org/10.1016/j.cub.2017.04.053

Waldenström, J., Bensch, S., Hasselquist, D., & Östman, Ö. (2004). A new nested polymerase chain reaction method very efficient in detecting *Plasmodium* and *Haemoproteus* infections from avian blood. *Journal of Parasitology*, 90(1), 191–194. https://doi.org/10.1645/GE-3221RN

Wang, S., Liu, C., Wilson, A. B., Zhao, N. A., Li, X., Zhu, W., … Li, Y. (2017). Pathogen richness and abundance predict patterns of adaptive major histocompatibility complex variation in insular amphibians. *Molecular Ecology*, 26, 4671–4685. https://doi.org/10.1111/mec.14242

Warren, W. C., Clayton, D. F., Ellegren, H., Arnold, A. P., Hillier, L. D. W., Künstner, A., … Wilson, R. K. (2010). The genome of a songbird. *Nature*, 464(7289), 757–762. https://doi.org/10.1038/nature08819

Wenzel, M. A., Douglas, A., James, M. C., Redpath, S. M., & Piertney, S. B. (2016). The role of parasite-driven selection in shaping landscape genomic structure in red grouse (*Lagopus lagopus scotica*). *Molecular Ecology*, 25(1), 324–341. https://doi.org/10.1111/mec.13473

Werling, D., Jann, O. C., Offord, V., Glass, E. J., & Coffey, T. J. (2009). Variation matters: TLR structure and species-specific pathogen recognition. *Trends in Immunology*, 30(3), 124–130. https://doi.org/10.1016/j.it.2008.12.001

Wlasiuk, G., & Nachman, M. W. (2010). Adaptation and constraint at toll-like receptors in primates. *Molecular Biology and Evolution*, 27(9), 2172–2186. https://doi.org/10.1093/molbev/msq104

Wong, S. H., Gochhait, S., Malhotra, D., Pettersson, F. H., Teo, Y. Y., Khor, C. C., … Vannberg, F. O. (2010). Leprosy and the adaptation of human toll-like receptor 1. *PLoS Pathogens*, 6(7), e1000979. https://doi.org/10.1371/journal.ppat.1000979

Yanoviak, S. P., Paredes, J. E. R., Lounibos, L. P., & Weaver, S. C. (2006). Deforestation alters phytotelm habitat availability and mosquito production in the Peruvian Amazon. *Ecological Applications*, 16(5), 1854–1864. https://doi.org/10.1890/1051-0761(2006)016[1854:DAPHAA]2.0.CO;2

Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, 1(1), 3–14. https://doi.org/10.1111/j.2041-210X.2009.00001.x